

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 04:30:35 ; Search time 4172.21 Seconds

(without alignments)
11406.588 Million cell updates/sec

Title: US-09-901-484A-179_COPY_3899_4996

Perfect score: 1098
Sequence: 1 gttctctgctcctgcgtctt.....agaacttaagttcctcgtgat 1098.

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_bcs:*
12: gb_ay:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ats:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_ay:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1098	100.0	56516	6	BD196412
2	1098	100.0	56520	6	BD196564
3	996	90.7	165799	2	AC100813
4	996	90.7	168136	9	AF287957
5	996	90.7	183485	2	AC022578
6	996	90.7	209718	2	AC091099
7	813	74.0	185834	2	AC009632
8	114	10.4	185834	2	AC009632
9	25	2.3	19087	6	AX345694
10	24	2.2	6782	6	AX345679
11	24	2.2	179514	10	AC134447
12	24	2.2	264864	2	AC135128
13	23	2.1	47	6	AR290033
14	23	2.1	47	6	AX114392
15	23	2.1	6070	6	AX281468
16	23	2.1	6070	6	AX346581
17	23	2.1	6070	6	AX348804
18	23	2.1	6070	6	AX458614
19	23	2.1	6074	6	AX323760
20	23	2.1	7238	6	AX345350
21	23	2.1	15121	6	AX344831
22	23	2.1	15649	6	AX348975
23	23	2.1	39403	3	CEB03A3
24	23	2.1	42740	2	AC090706
25	23	2.1	110394	9	AC003086
26	23	2.1	129109	9	AL390239
27	23	2.1	135862	9	AC025757
28	23	2.1	147874	2	AC011278
29	23	2.1	150438	9	AC099517
30	23	2.1	165965	2	AC012287
31	23	2.1	167122	2	AC053532
32	23	2.1	168216	2	AC013398
33	23	2.1	173445	9	AC078876
34	23	2.1	174617	9	AC105180
35	23	2.1	189464	2	AC112194
36	23	2.1	195909	9	AF165124
37	23	2.1	203131	9	AC009070
38	23	2.1	203982	9	AC067854
39	23	2.1	207040	9	AC091984
40	23	2.1	209836	9	AC091926
41	23	2.1	210736	2	AC092792
42	23	2.1	219223	2	BX649211
43	23	2.1	233219	2	AC113253
44	23	2.1	252420	3	AE014841
45	23	2.1	253149	2	AC107527

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
1	BD196412	Prostatic cancer gene.	BD196412	BD196412.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
2	BD196564	Prostatic cancer gene.	BD196564	BD196564.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
3	AC100813	Human DNA	AC100813	AC100813.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
4	AF287957	Human DNA	AF287957	AF287957.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
5	AC022578	Human DNA	AC022578	AC022578.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
6	AC091099	Human DNA	AC091099	AC091099.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
7	AC009632	Human DNA	AC009632	AC009632.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
8	AX345694	Human DNA	AX345694	AX345694.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
9	AX345679	Human DNA	AX345679	AX345679.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
10	AC134447	Human DNA	AC134447	AC134447.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
11	AC135128	Human DNA	AC135128	AC135128.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
12	AR290033	Human DNA	AR290033	AR290033.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
13	AX114392	Human DNA	AX114392	AX114392.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
14	AX281468	Human DNA	AX281468	AX281468.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
15	AX346581	Human DNA	AX346581	AX346581.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
16	AX348804	Human DNA	AX348804	AX348804.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
17	AX458614	Human DNA	AX458614	AX458614.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
18	AX323760	Human DNA	AX323760	AX323760.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
19	AX345350	Human DNA	AX345350	AX345350.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
20	AX344831	Human DNA	AX344831	AX344831.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
21	AX348975	Human DNA	AX348975	AX348975.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
22	CEB03A3	Human DNA	CEB03A3	CEB03A3.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
23	AC090706	Human DNA	AC090706	AC090706.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
24	AC003086	Human DNA	AC003086	AC003086.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
25	AL390239	Human DNA	AL390239	AL390239.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
26	AC025757	Human DNA	AC025757	AC025757.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
27	AC011278	Human DNA	AC011278	AC011278.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
28	AC099517	Human DNA	AC099517	AC099517.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
29	AC012287	Human DNA	AC012287	AC012287.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
30	AC053532	Human DNA	AC053532	AC053532.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
31	AC013398	Human DNA	AC013398	AC013398.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
32	AC078876	Human DNA	AC078876	AC078876.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
33	AC105180	Human DNA	AC105180	AC105180.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
34	AC112194	Human DNA	AC112194	AC112194.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
35	AF165124	Human DNA	AF165124	AF165124.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
36	AC009070	Human DNA	AC009070	AC009070.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
37	AC067854	Human DNA	AC067854	AC067854.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
38	AC091984	Human DNA	AC091984	AC091984.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
39	AC091926	Human DNA	AC091926	AC091926.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
40	AC092792	Human DNA	AC092792	AC092792.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
41	BX649211	Human DNA	BX649211	BX649211.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
42	AC113253	Human DNA	AC113253	AC113253.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
43	AE014841	Human DNA	AE014841	AE014841.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
44	AC107527	Human DNA	AC107527	AC107527.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 2002516657-A 11-JUN-2002;
45	AC107527	Human DNA	AC107527	AC107527.1	GI:33006182	JP 2002516657-A/1.	Homo sapiens (human)	1 (bases 1 to 56516)	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueteler, L.	Prostatic cancer gene	Patent: JP 20025166

[illegible]

FT	primer_bind	11622..	11639
FT	primer_bind	12018..	12037
FT	primer_bind	11930..	11947
FT	primer_bind	12339..	12358
FT	primer_bind	12915..	12932
FT	primer_bind	13317..	13334
FT	primer_bind	13216..	13233
FT	primer_bind	13617..	13636
FT	primer_bind	13547..	13564
FT	primer_bind	13962..	13981
FT	primer_bind	15994..	16011
FT	primer_bind	16463..	16480
FT	primer_bind	17304..	17324
FT	primer_bind	17814..	17832
FT	primer_bind	18008..	18027
FT	primer_bind	18423..	18442
FT	primer_bind	18699..	18716
FT	primer_bind	19164..	19182
FT	primer_bind	22589..	22609
FT	primer_bind	23111..	23129
FT	primer_bind	25098..	25118
FT	primer_bind	25657..	25674
FT	primer_bind	26537..	26557
FT	primer_bind	27022..	27040
FT	primer_bind	32262..	32281
FT	primer_bind	32823..	32841
FT	primer_bind	34215..	34233
FT	primer_bind	34624..	34644
FT	primer_bind	34473..	34491
FT	primer_bind	34916..	34936
FT	primer_bind	34702..	34722
FT	primer_bind	35182..	35200
FT	primer_bind	38581..	38611
FT	primer_bind	39711..	39791
FT	primer_bind	40531..	40549
FT	primer_bind	40932..	40950
FT	primer_bind	40629..	40649
FT	primer_bind	41058..	41078
FT	primer_bind	45729..	45746
FT	primer_bind	46186..	46203
FT	primer_bind	47879..	47896
FT	primer_bind	48217..	48236
FT	primer_bind	48902..	48922
FT	primer_bind	49331..	49351
FT	primer_bind	49830..	49848
FT	primer_bind	50271..	50290
FT	primer_bind	50172..	50189
FT	primer_bind	50573..	50591
FT	primer_bind	50541..	50560
FT	primer_bind	50940..	50959
FT	primer_bind	50555..	50572
FT	primer_bind	50964..	50983
FT	primer_bind	50774..	50792
FT	primer_bind	51183..	51202
FT	primer_bind	51146..	51165
FT	primer_bind	51479..	51496
FT	primer_bind	51593..	51610
FT	primer_bind	29734..	29744
FT	primer_bind	51167..	51185
FT	primer_bind	51667..	51687
FT	primer_bind	51387..	51403
FT	primer_bind	51826..	51843
FT	primer_bind	51772..	51789
FT	primer_bind	52199..	52217

Query Match 100.0%; Score 1098; DB 6; Length 56520;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1098; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCTGTCGCGCATTTTATTTATTTGATGAGGAGACAATATTTATTTCTGT 60
DB 3899 GTTCTGTCGCGCATTTTATTTATTTGATGAGGAGACAATATTTATTTCTGT 3958

QY	61	TTAGTAACAGAGGAGGTATTTGAAATTTATTTAGGGTCTTTTCTGACGTCTGGGTTCC	120
DB	3959	TTAGTAACAGAGGAGGTATTTGAAATTTATTTAGGGTCTTTTCTGACGTCTGGGTTCC	4018
QY	121	TGTGTACAAAGCTACCTTCAATATTTTATTTGTTTCTGTAAATTAATCAATAG	180
DB	4019	TGTGTACAAAGCTACCTTCAATATTTTATTTGTTTCTGTAAATTAATCAATAG	4078
QY	181	AGGAATTAATAGCATCTTCAACATAGACCCAAAGGAAAAAGATTATAGATGTTCC	240
DB	4079	AGGAATTAATAGCATCTTCAACATAGACCCAAAGGAAAAAGATTATAGATGTTCC	4138
QY	241	TGTACCTTATTTTATTTACCTGTGACTTTGTACATTAACCTTGTCACTAGATGTTTGA	300
DB	4139	TGTACCTTATTTTATTTACCTGTGACTTTGTACATTAACCTTGTCACTAGATGTTTGA	4198
QY	301	TTAAATTTTATTTAGCTTCTTTCTTTGTTTGTATAGACACTCTTTTCTGTAAATGTT	360
DB	4199	TTAAATTTTATTTAGCTTCTTTCTTTGTTTGTATAGACACTCTTTTCTGTAAATGTT	4258
QY	361	TTTATACGCTTGTGTTGCAAGGCTAGATGATTTCTTGTCTGTATTAAGTATTTGT	420
DB	4259	TTTATACGCTTGTGTTGCAAGGCTAGATGATTTCTTGTCTGTATTAAGTATTTGT	4318
QY	421	GACTCATTTCTGAAGGAGTTTATAGATTTTAAGAGTTATAGTTTAAATTAAGT	480
DB	4319	GACTCATTTCTGAAGGAGTTTATAGATTTTAAGAGTTATAGTTTAAATTAAGT	4378
QY	481	TTATTTATTTATATATTTATTAAGAGATTTTAAATTAATTTTATTAATGACATTT	540
DB	4379	TTATTTATTTATATATTTATTAAGAGATTTTAAATTAATTTTATTAATGACATTT	4438
QY	541	TTACACCTTTCACTCTGATTTTAAATTAAGTGTTCACAGTGTCTTGACAGAA	600
DB	4439	TTACACCTTTCACTCTGATTTTAAATTAAGTGTTCACAGTGTCTTGACAGAA	4498
QY	601	TATTTTCTTTTATCATAGATTTTATTTAGCTGAGAGGAGTATAGTCCATGATTT	660
DB	4499	TATTTTCTTTTATCATAGATTTTATTTAGCTGAGAGGAGTATAGTCCATGATTT	4558
QY	661	ATGATCTGTGCTGGCAGGTAACTGCTTCAACAAATTTAGTTGATTTTCTTGAT	720
DB	4559	ATGATCTGTGCTGGCAGGTAACTGCTTCAACAAATTTAGTTGATTTTCTTGAT	4618
QY	721	TCTGGGTAAATACCTTTTCTTCCCAAGTTTCACTACTTTATTTTATATGATCTGTA	780
DB	4619	TCTGGGTAAATACCTTTTCTTCCCAAGTTTCACTACTTTATTTTATATGATCTGTA	4678
QY	781	GATAGAGAAATTTTCACTGAGTGTGCTGTAATTTTCTTATTAACCTGTTATCTTT	840
DB	4679	GATAGAGAAATTTTCACTGAGTGTGCTGTAATTTTCTTATTAACCTGTTATCTTT	4738
QY	841	TAGGTCTTCCAGATCTCTCATTTGTGTAAGAACTCAATGGTACTTCTTCACTT	900
DB	4739	TAGGTCTTCCAGATCTCTCATTTGTGTAAGAACTCAATGGTACTTCTTCACTT	4798
QY	901	TATTTCTTTAGAAATAGTAATTAAGATTTTATTAAGCTTTTATTTATTTCACTTGA	960
DB	4799	TATTTCTTTAGAAATAGTAATTAAGATTTTATTAAGCTTTTATTTATTTCACTTGA	4858
QY	961	GACTATTTGAAATCCAGTTAAGTCTCTCACTGCTGTGAGAGGATTAATCAAGTACT	1020
DB	4859	GACTATTTGAAATCCAGTTAAGTCTCTCACTGCTGTGAGAGGATTAATCAAGTACT	4918
QY	1021	GTTTACTTCTGTGCTGCCAAAACAGATCACTCAATAGAGGCTTAAATTAATAG	1080
DB	4919	GTTTACTTCTGTGCTGCCAAAACAGATCACTCAATAGAGGCTTAAATTAATAG	4978
QY	1081	AACTTAAGTTCTGTGAT 1098	
DB	4979	AACTTAAGTTCTGTGAT 4996	

QY 601 TATTTCTTTTACATAGAAATTTTAAAGCTGAAGAGAGAGTAAAGGTCCATGAGATT 660
Db 51474 TATTTCTTTTACATAGAAATTTTAAAGCTGAAGAGAGAGTAAAGGTCCATGAGATT 51415
QY 661 ATGATCTGTGCTTGAGCAGGTAACCTGCTCCACAAATTTAGTTGATTTTCTTGAT 720
Db 51414 ATGATCTGTGCTTGAGCAGGTAACCTGCTCCACAAATTTAGTTGATTTTCTTGAT 51355
QY 721 TCTGGGTAATACCTTTTCTTCCCAAGTTTACACTTATTTTCAATATGATCTCTGA 780
Db 51354 TCTGGGTAATACCTTTTCTTCCCAAGTTTACACTTATTTTCAATATGATCTCTGA 51295
QY 781 GATPAGAGAAATTTTACAGTCAGTCGCTAAATTTGTCCTATATGCTTATCTCTT 840
Db 51294 GATPAGAGAAATTTTACAGTCAGTCGCTAAATTTGTCCTATATGCTTATCTCTT 51235
QY 841 TAGGTCTCTCCAGAAATCTCTCATTTGGTACTGAAGTCAATGGTACTTTTCCACCAAT 900
Db 51234 TAGGTCTCTCCAGAAATCTCTCATTTGGTACTGAAGTCAATGGTACTTTTCCACCAAT 51175
QY 901 TATTTCTTTAGAAATAGTAATTAAGATTTTATAGCTTTTATATTTTCACTATTTGA 960
Db 51174 TATTTCTTTAGAAATAGTAATTAAGATTTTATAGCTTTTATATTTTCACTATTTGA 51115
QY 961 GACTATTTGAAATCCAGTTAAGTCTCTACTGCTGTGAGAGGCAATGATCAAGTACT 1020
Db 51114 GACTATTTGAAATCCAGTTAAGTCTCTACTGCTGTGAGAGGCAATGATCAAGTACT 51055
QY 1021 GTGTACTTCTCTGTGCTGCCAAACAGATCACTCAACTAAGCGGCTTAATAATPAIN 1080
Db 51054 GTGTACTTCTCTGTGCTGCCAAACAGATCACTCAACTAAGCGGCTTAATAATPAIN 50995
QY 1081 AACTTAAGTTCTCGTAT 1098
Db 50994 AACTTAAGTTCTCGTAT 50977

RESULT 4
AF287957 168136 bp DNA linear PRI 08-JUL-2003
LOCUS Homo sapiens chromosome 8 clone CTD-2541M15 map p22-p21, complete
DEFINITION
sequence.
ACCESSION AF287957
VERSION AF287957.6 GI:32469520
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 168136)
Reichwald,K., Baumgart,C., Blechschmidt,K., Dettre,M., Jahn,N.,
Wenzel,U., Polley,A., Schilhabel,M.B., Taudien,S., Wen,G. and
Rosenthal,A.
Chromosome 8 genomic sequence
2 (bases 1 to 168136)
Unpublished
3 (bases 1 to 168136)
Direct Submission
Submitted (14-JUL-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
Reichwald,K. and Platzer,M.
Direct Submission
Submitted (23-AUG-2001) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
4 (bases 1 to 168136)
Lagemann,D. and Platzer,M.
Direct Submission
Submitted (29-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
5 (bases 1 to 168136)
Lagemann,D. and Platzer,M.

TITLE Direct Submission
JOURNAL Submitted (03-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE 6 (bases 1 to 168136)
AUTHORS Lagemann,D. and Platzer,M.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 7 (bases 1 to 168136)
AUTHORS Lagemann,D. and Platzer,M.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On Jul 8, 2003 this sequence version replaced gi:21717104.

Center: Institute of Molecular Biotechnology

Center code: IMB
Web site: <http://genome.imb-jena.de/>
Contact: gsccj-submit@genome.imb-jena.de

Project Information

Center project name: H351
Center clone name: CTD-2541M15

Summary Statistics

Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 167707 bases at least Q40
Consensus quality: 168123 bases at least Q30
Consensus quality: 168136 bases at least Q20
Quality coverage: 6.18x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one pUC18 subclone, and the
assembly was confirmed by restriction digest.

Neighboring sequence information:
This clone is overlapped by GS1-24F4, GS1-24F4.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

Location/Qualifiers
1. 168136
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="p22-p21"
/clone="CTD-2541M15"
165266..168136
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="GS1-24F4"
/note="overlapping clone, overlapping clone"
6418
/note="low quality region"
7442..7645
/note="single stranded/single chemistry region"
28243..28507
/note="single stranded/single chemistry region"
28499

misc_feature	/note="low quality region"	28501
misc_feature	/note="low quality region"	34532.34643
misc_feature	/note="single stranded/single chemistry region"	35997.36115
misc_feature	/note="single stranded/single chemistry region"	37530
misc_feature	/note="low quality region"	38752.38820
misc_feature	/note="single stranded/single chemistry region"	38995.39017
misc_feature	/note="single stranded/single chemistry region"	44702.44825
misc_feature	/note="single stranded/single chemistry region"	45312.45406
misc_feature	/note="single stranded/single chemistry region"	47601.47995
misc_feature	/note="single stranded/single chemistry region"	59027.59136
misc_feature	/note="single stranded/single chemistry region"	6119
misc_feature	/note="low quality region"	62590.62726
misc_feature	/note="single stranded/single chemistry region"	66851.66899
misc_feature	/note="single stranded/single chemistry region"	69911.70188
misc_feature	/note="single stranded/single chemistry region"	75977.76621
misc_feature	/note="single stranded/single chemistry region"	78279.78386
misc_feature	/note="single stranded/single chemistry region"	78675
misc_feature	/note="low quality region"	81011.81279
misc_feature	/note="single stranded/single chemistry region"	81070.81071
misc_feature	/note="low quality region"	81743.82001
misc_feature	/note="single stranded/single chemistry region"	81872
misc_feature	/note="low quality region"	84434.85028
misc_feature	/note="pcr product sequence only"	84932.85028
misc_feature	/note="single stranded/single chemistry region"	86073.86997
misc_feature	/note="single stranded/single chemistry region"	86609
misc_feature	/note="low quality region"	89636.89783
misc_feature	/note="single stranded/single chemistry region"	90439.90540
misc_feature	/note="single stranded/single chemistry region"	95593.95693
misc_feature	/note="single stranded/single chemistry region"	98827.98848
misc_feature	/note="single stranded/single chemistry region"	98860
misc_feature	/note="low quality region"	99495.99715
misc_feature	/note="single stranded/single chemistry region"	10416.105145
misc_feature	/note="single stranded/single chemistry region"	105357.105404
misc_feature	/note="single stranded/single chemistry region"	105600.106813
misc_feature	/note="pcr product sequence only"	112247.112576
misc_feature	/note="single stranded/single chemistry region"	117025.117051
misc_feature	/note="single stranded/single chemistry region"	

	misc_feature	118344. 118414	/note="single stranded/single chemistry region"
	misc_feature	141144. 141340	/note="single stranded/single chemistry region"
	misc_feature	141695. 141700	/note="single stranded/single chemistry region"
	misc_feature	142575. 142734	/note="single stranded/single chemistry region"
	misc_feature	145776. 145927	/note="single stranded/single chemistry region"
	misc_feature	146784. 146859	/note="single stranded/single chemistry region"
	misc_feature	150458. 150522	/note="single stranded/single chemistry region"
	misc_feature	160235. 161053	/note="single stranded/single chemistry region"
	misc_feature	163686. 163687	/note="single stranded/single chemistry region"
	misc_feature	165389. 165603	/note="low quality region"
	misc_feature		/note="single stranded/single chemistry region"
ORIGIN			
Query Match	90.7%;	Score 996;	DB 9; Length 168136;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 1096;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
QY	1	GTTCCTGTCCTCGCATTTTATTTTATTTGATGAGGGGCAATAATATTTTCTGT	60
DB	116065	GTTTCCTGTCCTCGCATTTTATTTTATTTGATGAGGGGCAATAATATTTTCTGT	116124
QY	61	TTAGTACAGAGCAGGATTTTGAATTAATAGGTCCTTTTCTGACGTCGGGTTCC	120
DB	116125	TTAGTACAGAGCAGGATTTTGAATTAATAGGTCCTTTTCTGCAATCTGGGTTCC	116184
QY	121	TGTGACACCAAGCATCCTTCAATTTTATTTGTTCTGTGAAGATTAATCAATAG	180
DB	116185	TGTGACACCAAGCATCCTTCAATTTTATTTGTTCTGTGAAGATTAATCAATAG	116244
QY	181	AGGATAAATAGCTATCTTCAACACATAAGACCCAAAGAAAAAGATTATAGATGTT	240
DB	116245	AGGATAAATAGCTATCTTCAACACATAAGACCCAAAGAAAAAGATTATAGATGTT	116304
QY	241	TGTCACTTAATTTTACCTGTGACTTTGACATTAACTTTGTCACTGAGATGTTTGA	300
DB	116305	TGTCACTTAATTTTACCTGTGACTTTGACATTAACTTTGTCACTGAGATGTTTGA	116364
QY	301	TTAAATTTTATAGCTTGCCTTTCTGTTTGTAGACCTCTTTTCTGGAATGTT	360
DB	116365	TTAAATTTTATAGCTTGCCTTTCTGTTTGTAGACCTCTTTTCTGGAATGTT	116424
QY	361	TTTATCAGCTTTCGTTTCCAAAGGCTAGTGAATTCCTGTGTCGTATAAAGTATGTT	420
DB	116425	TTTATCAGCTTTCGTTTCCAAAGGCTAGTGAATTCCTGTGTCGTATAAAGTATGTT	116484
QY	421	GACTCATTTCTGAAGGAGTTTATGTAATTTAAGGTTATAAGTTTATAATAAAGGT	480
DB	116485	GACTCATTTCTGAAGGAGTTTATGTAATTTAAGGTTATAAGTTTATAATAAAGGT	116544
QY	481	TTATTAATTTATATATTAATAAGGCACTTTAAATATAATTTTATAATGACATTT	540
DB	116545	TTATTAATTTATATATTAATAAGGCACTTTAAATATAATTTTATAATGACATTT	116604
QY	541	TTACACCTTCAACTCTAGGTTTAAATAAATAAGTGTTCACAGTATGTTCTTGCAAGAA	600
DB	116605	TTACACCTTCAACTCTAGGTTTAAATAAATAAGTGTTCACAGTATGTTCTTGCAAGAA	116664
QY	601	TATTTTCTTTTACATAGAAATTTTAACTGAAGAGAGTAGTAGTTCATGAGATTT	660
DB	116665	TATTTTCTTTTACATAGAAATTTTAACTGAAGAGAGTAGTAGTTCATGAGATTT	116724
QY	661	ATGATCTGTGCTGGCAGGTAAACCTGCTCCAAAGAAATTAGTGGATTTTCTTGAT	720
DB	116725	ATGATCTGTGCTGGCAGGTAAACCTGCTCCAAAGAAATTAGTGGATTTTCTTGAT	116784

QY 721 TCTGGGTAATACCTTTTCTTCCCAAGTTTCACTTATTTTCAATATGATCTCGA 780
 Db 116785 TCTGGGTAATACCTTTTCTTCCCAATTTTCACTTATTTTCAATATGATCTCGA 116844
 QY 761 GATTAGAAATATTTTCACTGAGTGTGCTGAAATATGTTTCTTATTAAGTATCTCTT 840
 Db 116845 GATTAGAAATATTTTCACTGAGTGTGCTGAAATATGTTTCTTATTAAGTATCTCTT 116904
 QY 841 TAGGTCTTCCAGAAATCTCTCACTGTTGTTGTAAGTCAAAATGGTACTTTCTTCACTT 900
 Db 116905 TAGGTCTTCCAGAAATCTCTCACTGTTGTTGTAAGTCAAAATGGTACTTTCTTCACTT 116964
 QY 901 TATTTCTTGAATAGTAATAGTAATAGTAATAGTAATAGTAATAGTAATAGTAATAG 960
 Db 116965 TATTTCTTGAATAGTAATAGTAATAGTAATAGTAATAGTAATAGTAATAGTAATAG 117024
 QY 961 GACTATTTGAATATCCAGTAATAGTCTCTCACTGTTGTTGTAAGTCAAAATGGTACTT 1020
 Db 117025 GACTATTTGAATATCCAGTAATAGTCTCTCACTGTTGTTGTAAGTCAAAATGGTACTT 117084
 QY 1021 GTGTACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 117085 GTGTACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 117144
 QY 1081 AACTTAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098
 Db 117145 AACTTAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 117162

RESULT 5
 AC022578 183485 bp DNA linear HNG 16-MAR-2001
 LOCUS Homo sapiens clone RP11-15C20, WORKING DRAFT SEQUENCE, 11 unordered
 DEFINITION pieces.
 AC022578
 VERSION AC022578.5 GI:13357410
 HGVS: HNGS PHASE1: HNGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 183485)
 Birren,B., Linton,L., Nussbaum,C. and Lander,E.
 Homo sapiens, clone RP11-15C20
 Unpublished
 2 (bases 1 to 183485)
 Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckert,R., Beda,F.,
 Bogunlavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearlano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J.,
 Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karas,A., Klein,U.,
 Landers,T., Lehoczeky,J., Levine,R., Lien,C., Liu,G., Locke,K.,
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
 Mcpheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
 Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
 Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Mar 16, 2001 this sequence version replaced gi:1231816.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: U3441
 Center clone name: 15_C_20
 Summary Statistics
 Sequencing vector: M13; M7815; 4% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 179658 bases at least Q40
 Consensus quality: 181356 bases at least Q30
 Consensus quality: 182011 bases at least Q20
 Insert size: 170000; agarose-fp
 Insert size: 182485; sum-of-ctigs
 Quality coverage: 7.0 in Q20 bases; agarose-fp
 Quality coverage: 6.5 in Q20 bases.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 11 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

Contig	Start	End	Length	Quality
1	1387	contig of 1387 bp	in length	1387
1388	1487	gap of 100 bp	in length	1487
1488	2538	contig of 1051 bp	in length	2538
2539	2638	gap of 100 bp	in length	2638
2639	4524	contig of 1886 bp	in length	4524
4525	4624	gap of 100 bp	in length	4624
4625	7391	contig of 2767 bp	in length	7391
7392	7492	gap of 100 bp	in length	7492
7493	9933	contig of 2442 bp	in length	9933
9934	10033	gap of 100 bp	in length	10033
10034	14280	contig of 4247 bp	in length	14280
14281	14380	gap of 100 bp	in length	14380
14381	25807	contig of 11427 bp	in length	25807
25808	25907	gap of 100 bp	in length	25907
25908	43775	contig of 17868 bp	in length	43775
43776	43875	gap of 100 bp	in length	43875
43876	67057	contig of 23182 bp	in length	67057
67058	67157	gap of 100 bp	in length	67157
67158	126378	contig of 59121 bp	in length	126378
126379	126379	gap of 100 bp	in length	126379
126379	183485	contig of 57107 bp	in length	183485

Location/Qualifiers
 1. 183485
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-15C20"
 /clone_lib="RPCI-11 Human Male BAC"
 1. 1387
 /note="assembly_fragment"
 1488. 2538
 /note="assembly_fragment"
 2639. 4524
 /note="assembly_fragment"
 4625. 7391
 /note="assembly_fragment"
 7492. 9933
 /note="assembly_fragment"
 10034. 14280
 /note="assembly_fragment"
 14381. 25807
 /note="assembly_fragment"
 25908. 43775
 /note="assembly_fragment"
 43876. 67057
 /note="assembly_fragment"
 67158. 126378

TITLE
JOURNAL
COMMENT

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:15290844.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L12159

Center clone name: 782_N_1

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 206063 bases at least Q40

Consensus quality: 207712 bases at least Q30

Insert size: 205000; agarose-fp

Insert size: 208818; sum-of-contigs

Quality coverage: 8.8 in Q20 bases; agarose-fp

Quality coverage: 8.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 74411: contig of 74411 bp in length
* 74412 74511: gap of 100 bp
* 74512 74749: contig of 238 bp in length
* 74750 74849: gap of 100 bp
* 74850 76857: contig of 2008 bp in length
* 76858 76957: gap of 100 bp
* 76958 80208: contig of 3251 bp in length
* 80209 80308: gap of 100 bp
* 80309 88877: contig of 8569 bp in length
* 88878 88977: gap of 100 bp
* 88978 104129: contig of 15152 bp in length
* 104130 104229: gap of 100 bp
* 104230 120104: contig of 15875 bp in length
* 120105 120204: gap of 100 bp
* 120205 152189: contig of 31985 bp in length
* 152190 152289: gap of 100 bp
* 152290 208329: contig of 56040 bp in length
* 208330 208429: gap of 100 bp
* 208430 209718: contig of 1289 bp in length.

FEATURES

SOURCE

1. 209718

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-782N1"

/clone_lib="RP11 Human Male BAC"

1. 74411

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

74512..74749

/note="assembly_fragment"

74850..76857

/note="assembly_fragment"

ORIGIN

Query Match 90.7%; Score 996; DB 2; Length 209718;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1096; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 GTTCTCTGTCCTCCGATTTTATTTATGTATGAGGAGGACAAATATTTCTCTG 60
Db 130520 GTTCTCTGTCCTCCGATTTTATTTATGTATGAGGAGGACAAATATTTCTCTG 130579
61 TTAGTAACAGACGAGGATTTGAATTTATAGGCTCTTTCTGACAGTCTGGGTTCC 120
Db 130580 TTAGTAACAGACGAGGATTTGAATTTATAGGCTCTTTCTGACAGTCTGGGTTCC 130639
121 TGTGTACCAAGGCTACCTTCAATTTTATTTATTTCTGTTAAGATTAAATCAATAG 180
Db 130640 TGTGTACCAAGGCTACCTTCAATTTTATTTATTTCTGTTAAGATTAAATCAATAG 130699
181 AGGAATTAATGCAATCTTCAATTAAGACCAAGGAAAGAAATTAATGATGTC 240
Db 130700 AGGAATTAATGCAATCTTCAATTAAGACCAAGGAAAGAAATTAATGATGTC 130759
241 TGTACCTTATTTTACCTGCTGACCTTGTACCTTAATCTTGTCACTGATGTTTGA 300
Db 130760 TGTACCTTATTTTACCTGCTGACCTTGTACCTTAATCTTGTCACTGATGTTTGA 130819
301 TTTAAATTTTACCTGCTTCTGTTTGTGTTTGTAGACACTCTTTTCTTGAATGTT 360
Db 130820 TTTAAATTTTACCTGCTTCTGTTTGTGTTTGTAGACACTCTTTTCTTGAATGTT 130879
361 TTTATGACCTTCTGTTTCAAGGCTAGATGATCTCTGTTGTAATGAATGTT 420
Db 130880 TTTATGACCTTCTGTTTCAAGGCTAGATGATCTCTGTTGTAATGAATGTT 130939
421 GACTCATTTCTGAAGGAGTTTATGATTTTAAAGGTTATAGTTTAAATTAAGT 480
Db 130940 GACTCATTTCTGAAGGAGTTTATGATTTTAAAGGTTATAGTTTAAATTAAGT 130999
481 TTATTAATTTATATATTAATTAAGGCAATTTAAATTAATTTTAAATGACATT 540
Db 131000 TTATTAATTTATATATTAATTAAGGCAATTTAAATTAATTTTAAATGACATT 131059
541 TTACACCTTTCAACCTTGAAGTTTAAATTAAGGTTCAAGTCTTCAAGAGAA 600
Db 131060 TTACACCTTTCAACCTTGAAGTTTAAATTAAGGTTCAAGTCTTCAAGAGAA 131119
601 TATTTCTTTTACATAGATTTTAAAGTGAAGAGATAGTAGTCCATGAGATT 660
Db 131120 TATTTCTTTTACATAGATTTTAAAGTGAAGAGATAGTAGTCCATGAGATT 131179
661 ATGATCTGCTTGGCAGTAAACCTGCTTCAACAAATTTAGTTGATTTTCTTGAT 720
Db 131180 ATGATCTGCTTGGCAGTAAACCTGCTTCAACAAATTTAGTTGATTTTCTTGAT 131239
721 TCTGGTAAATACCTTTTCTCCCGAGTTCACTACTTATTTTCAATGATCTCGA 780
Db 131240 TCTGGTAAATACCTTTTCTCCCGAGTTCACTACTTATTTTCAATGATCTCGA 131299

QY 781 GATAGAGAAATATTTCAGTGCAGTGCCTAAATTTGTTCTTATPACTGTTATCCCTT 840
|||
Db 131300 GATAGAGAAATATTTCAGTGCAGTGCCTAAATTTGTTCTTATPACTGTTATCCCTT 131359
QY 841 TAGGTCCCTTCAGAAATCTCTGATTTGTAAGTAACTCAATGGTACTTTCTTCACCAT 900
131360 TAGGTCCCTTCAGAAATCTCTGATTTGTAAGTAACTCAATGGTACTTTCTTCACCAT 131419
QY 901 TATTTCTTAAATTAAGTAATTAAGAAATTTTATAGCTTTTATTTATTTTACAGTAATTTGA 960
131420 TATTTCTTAAATTAAGTAATTAAGAAATTTTATAGCTTTTATTTATTTTACAGTAATTTGA 131479
QY 961 GACTATTGAATTAAGTAATTAAGTAATTTTATAGCTTTTATTTATTTTACAGTAATTTGA 1020
131480 GACTATTGAATTAAGTAATTAAGTAATTTTATAGCTTTTATTTATTTTACAGTAATTTGA 131539
QY 1021 GTGTACTTTCTGCTGTGCTGCAAAAAGATGACCTCAAACTAAGGCGCTTAAATTAATAG 1080
131540 GTGTACTTTCTGCTGTGCTGCAAAAAGATGACCTCAAACTAAGGCGCTTAAATTAATAG 131599
QY 1081 AACTTAAGTTCTCGTAT 1098
131600 AACTTAAGTTCTCGTAT 131617
Db

RESULT 7
AC009632 185834 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 8 clone RP11-397K22 map 8, *** SEQUENCING
DEFINITION IN PROGRESS ***, 48 unordered pieces.
AC009632
AC009632.3 GI:8954116
VERSION HTG, HTGS, PHASE1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Unpublished
JOURNAL 2 (bases 1 to 185834)
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,
Cacile, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArrellano, K., Depayre, B., Devon, K., Dewar, K.,
Domelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
Karats, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P.,
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meidrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Nijloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Teefaye, S., Tornella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
DIRECT SUBMISSION
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 7, 2000 this sequence version replaced gi:5932606.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL
COMMENT
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L2099
Center clone name: 397_K_22

* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1361: contig of 1361 bp in length
1362 1461: gap of 100 bp
1462 3202: contig of 1741 bp in length
3203 3302: gap of 100 bp
3303 4619: contig of 1317 bp in length
4620 4719: gap of 100 bp
4720 5821: contig of 1102 bp in length
5822 5921: gap of 100 bp
5922 6419: contig of 498 bp in length
6420 6520: gap of 100 bp
6520 8242: contig of 1723 bp in length
8243 8343: gap of 100 bp
8343 9703: contig of 1361 bp in length
9704 9803: gap of 100 bp
9804 10936: contig of 1133 bp in length
10937 11037: gap of 100 bp
11037 12339: contig of 1303 bp in length
12340 12439: gap of 100 bp
12440 13923: contig of 1484 bp in length
13924 14023: gap of 100 bp
14024 15372: contig of 1349 bp in length
15373 15472: gap of 100 bp
15473 16957: contig of 1485 bp in length
16958 17057: gap of 100 bp
17058 19291: contig of 2234 bp in length
19292 19391: gap of 100 bp
19392 21442: contig of 2051 bp in length
21443 21542: gap of 100 bp
21543 23482: contig of 1940 bp in length
23483 23582: gap of 100 bp
23583 25557: contig of 1975 bp in length
25558 26557: gap of 100 bp
26558 28144: contig of 2487 bp in length
28145 28244: gap of 100 bp
28245 30739: contig of 2495 bp in length
30740 30839: gap of 100 bp
30840 33278: contig of 2439 bp in length
33279 33378: gap of 100 bp
33379 35479: contig of 2101 bp in length
35480 35579: gap of 100 bp
35580 38046: contig of 2467 bp in length
38047 38146: gap of 100 bp
38147 40591: contig of 2445 bp in length
40592 40691: gap of 100 bp
40692 43194: contig of 2503 bp in length
43195 43294: gap of 100 bp
43295 46358: contig of 3064 bp in length
46359 46458: gap of 100 bp
46459 49611: contig of 3153 bp in length
49612 49711: gap of 100 bp
49712 53161: contig of 3450 bp in length
53162 53261: gap of 100 bp
53262 55787: contig of 2526 bp in length
55788 55887: gap of 100 bp
55889 58727: contig of 2840 bp in length
58728 58827: gap of 100 bp
58828 61970: contig of 3143 bp in length
61971 62070: gap of 100 bp
62071 65666: contig of 3596 bp in length
65667 70066: gap of 100 bp
70067 70106: contig of 4240 bp in length
70107 73614: gap of 100 bp
73615 73714: contig of 3508 bp in length

*	73715	78129:	contig of 4415 bp in length
*	78130	78229:	gap of 100 bp
*	78230	80922:	contig of 2673 bp in length
*	80903	81002:	gap of 100 bp
*	81003	84942:	contig of 3940 bp in length
*	84943	85042:	gap of 100 bp
*	85043	90417:	contig of 5375 bp in length
*	90418	90517:	gap of 100 bp
*	90518	95577:	contig of 5080 bp in length
*	95598	95637:	gap of 100 bp
*	95637	101500:	contig of 5803 bp in length
*	95959	101501:	gap of 100 bp
*	101501	101600:	gap of 100 bp
*	101601	108688:	contig of 7088 bp in length
*	108689	108788:	gap of 100 bp
*	108789	115623:	contig of 6835 bp in length
*	115624	115723:	gap of 100 bp
*	115724	120744:	contig of 5021 bp in length
*	120745	120844:	gap of 100 bp
*	120845	128016:	contig of 7172 bp in length
*	128017	128116:	gap of 100 bp
*	128117	133590:	contig of 5574 bp in length
*	133591	133790:	gap of 100 bp
*	133791	141605:	contig of 7815 bp in length
*	141606	141705:	gap of 100 bp
*	141706	150654:	contig of 8949 bp in length
*	150655	150754:	gap of 100 bp
*	150755	158712:	contig of 7958 bp in length
*	158713	158812:	gap of 100 bp
*	158813	166914:	contig of 10104 bp in length
*	166917	168016:	gap of 100 bp
*	169017	168534:	contig of 16818 bp in length.

FEATURES	Location/Qualifiers
source	1. .185834

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/Chromosome="8"
/map="8"
/clone="RP11-397K22"
/clone_1fb="RP11-11 Human Male BAC"
misc_feature
1..1361
/note="assembly_fragment"
1462..3202
/note="assembly_fragment"
3303..4619
/note="assembly_fragment"
4720..5821
/note="assembly_fragment"
5922..6419
/note="assembly_fragment"
clone_end:"7"
vector_side:"right"
6520..8242
/note="assembly_fragment"
8343..9703
/note="assembly_fragment"
9804..10936
/note="assembly_fragment"
11037..12339
/note="assembly_fragment"
12440..13923
/note="assembly_fragment"
14024..15372
/note="assembly_fragment"
15473..16957
/note="assembly_fragment"
17058..19291
/note="assembly_fragment"
19392..21442
/note="assembly_fragment"
21543..22482
/note="assembly_fragment"
23583..25557
misc_feature

```

```
misc_feature  /note="assembly_fragment"  
25558. 2014  
note="assembly_fragment"  
misc_feature  2845. 30739  
/note="assembly_fragment"  
30840. 33278  
note="assembly_fragment"  
misc_feature  33379. 35479  
/note="assembly_fragment"
```

Query Match	74.0%;	Score 813;	DB 2;	Length 185834;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 863;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

OY	235	ATGTCCTGCACTTATTTTAACTTGACCTTGACCTTGTACCAATTACCTTGGCACCTGGAGAT	294
Db	43295	ATGTCCTGCACTTATTTTAACTTGACCTTGACCTTGTACCAATTACCTTGGCACCTGGAGAT	43354
OY	295	TTTGTATTTAAATTTTAACTTGCTTGGCTTTCTGTGTTTGTATGAGCACACTCTTTTTCCTGA	354
Db	43355	TTTGTATTTAAATTTTAACTTGCTTGGCTTTCTGTGTTTGTATGAGCACACTCTTTTTCCTGA	43414
OY	355	ATTGTTTTTATCAGCTTTCGTTTGGCAAGCTAGTGAATTCCTGTCCTGTATTAAGT	414
Db	43415	ATTGTTTTTATCAGCTTTCGTTTGGCAAGCTAGTGAATTCCTGTCCTGTATTAAGT	43474
OY	415	ATTGTTGACCTATTTCTGAAAGGAGTTTATTAATTTAAGAGTTTAAAGTTTAAATA	474
Db	43475	ATTGTTGACCTATTTCTGAAAGGAGTTTATTAATTTAAGAGTTTAAAGTTTAAATA	43534
OY	475	AAAGTTTATTTATTTATTTATTTATTTAAAGGACATTTTAAATTTTAAAAATTTTAAATG	534
Db	43535	AAAGTTTATTTATTTATTTATTTATTTAAAGGACATTTTAAATTTTAAAAATTTTAAATG	43594
OY	535	ACATTTTATACCTTTCAACTCTAGTTTAAATAAATAGTGGTTCACAGTAGTCTTGCA	594
Db	43595	ACATTTTATACCTTTCAACTCTAGTTTAAATAAATAGTGGTTCACAGTAGTCTTGCA	43654
OY	595	GAGAATATTTTCTTTTACATGAATTTTAAAGCTGAAGAGAGTAGTAGTAGGTCATG	654
Db	43655	GAGAATATTTTCTTTTACATGAATTTTAAAGCTGAAGAGAGTAGTAGTAGGTCATG	43714
OY	655	AGATTTATATATCTGTCCTTGCAAGGTAACCTGCTTCCAAACAAATTTAGTGGATTTTTC	714
Db	43715	AGATTTATATATCTGTCCTTGCAAGGTAACCTGCTTCCAAACAAATTTAGTGGATTTTTC	43774
OY	715	TTGAGATTCGGGTAATACTTTTTCCTCCCAAGTTTCACTATTATTTTCATATGAT	774
Db	43775	TTGAGATTCGGGTAATACTTTTTCCTCCCAAGTTTCACTATTATTTTCATATGAT	43834
OY	775	CTCTGAGATGAGAAATATTTTACAGCAGTGTGCTGCTAAATTTGTTCTTATTAAGCTGTTA	834
Db	43835	CTCTGAGATGAGAAATATTTTACAGCAGTGTGCTGCTAAATTTGTTCTTATTAAGCTGTTA	43894
OY	835	TTCCTTTAGGTCCTTCCAGATCTCTCATTTGTAAGTAACTCAATGGGTACTTTCTTC	894
Db	43895	TTCCTTTAGGTCCTTCCAGATCTCTCATTTGTAAGTAACTCAATGGGTACTTTCTTC	43954
OY	895	ACCAATTATTTCTTTTGAATTAAGTAATGAATTTTAAAGTTTATTAATTTTAAAGCTGA	954
Db	43955	ACCAATTATTTCTTTTGAATTAAGTAATGAATTTTAAAGTTTATTAATTTTAAAGCTGA	44014
OY	955	ATTTGAGACATTTGAATAATCAGTTAAGTCTCTTAAGTGTGTTGAAGGCAATGATTCOA	1014
Db	44015	ATTTGAGACATTTGAATAATCAGTTAAGTCTCTTAAGTGTGTTGAAGGCAATGATTCOA	44074
OY	1015	GTAACCTGTGTTACTTTCCTGTGTGCTGCCAAACAGATCACCTTAATAAGCGGCTTAAAT	1074
Db	44075	GTAACCTGTGTTACTTTCCTGTGTGCTGCCAAACAGATCACCTTAATAAGCGGCTTAAAT	44134
OY	1075	TAATAGAATTTAAGTCTCGTGAT	1098
Db	44135	TAATAGAATTTAAGTCTCGTGAT	44158

KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 179514)
Wilson, R.K.
TITLE
The sequence of Mus musculus clone
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 179514)
McPherson, J.D. and Waterston, R.H.
AUTHORS
Direct Submission
JOURNAL
Submitted (26-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 179514)
Wilson, R.K.
REFERENCE
Direct Submission
AUTHORS
Submitted (12-DEC-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
On Dec 12, 2003 this sequence version replaced gi:23322797.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
Project information
Center project name: M_BB0186D01

FEATURES
source
Location/Qualifiers
1. 179514
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="18"
/clone="RP24-186D1"

ORIGIN
Query Match 2.2% Score 24; DB 10; Length 179514;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
480 TTTATTAATTTATATATTAAG 503
Db 141070 TTTATTAATTTATATTAAG 141047

RESULT 12
AC135128/c
AC135128 264864 bp DNA linear HTG 20-NOV-2002
LOCUS
DEFINITION
Rattus norvegicus clone CH230-249P17, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
AC135128
AC135128.2 GI:25138667
HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 264864)
Munry, D., Marie, Metzger, M., Lee, A., Ammon, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Behnmed, F., Bialwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deremo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

TITLE
Unpublished
JOURNAL
2 (bases 1 to 264864)
REFERENCE
Rat Genome Sequencing Consortium.
AUTHORS
Direct Submission
JOURNAL
Submitted (07-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 264864)
Rat Genome Sequencing Consortium.
REFERENCE
Rat Genome Sequencing Consortium.
AUTHORS
Direct Submission
JOURNAL
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23507717.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: KC2U
Center clone name: CH230-249P17
----- Summary Statistics -----

Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geier, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulvik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewtas, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisse, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Miosavljevic, A., Miner, G., Mirza, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokedeme, O., Okunolu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poldexter, A., Popov, D., Primus, B., Pu, L., Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinkey, A., Trejos, Z., Usmani, K., Valas, R., Vera, Y., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhu, D., von Weinstock, G., and Gibbs, R.A.

Assembly program: Phrap; version 0.990329
Consensus quality: 239879 bases at least Q40
Consensus quality: 242628 bases at least Q20
Consensus quality: 244265 bases at least Q20
Estimated insert size: 242071; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 263304: contig of 263304 bp in length
* 263305 263404: gap of unknown length
* 263405 264864: contig of 1460 bp in length.
Location/Qualifiers

FEATURES
source
1..264864
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-249F17"
1..1215
/note="wgs_end_extension
clone_end:Sp6"
4648..5394
/note="clone boundary
clone_end:Sp6
site:"
end_sequence:BZ213261"
100740..102557
/note="wgs_contig"
complement(259844..260595)
/note="clone boundary
clone_end:T7
site:
end_sequence:BZ213258"
261463..263304
/note="wgs_end_extension
clone_end:T7"

ORIGIN

Query Match 2.2%; Score 24; DB 2; Length 264864;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 599 AATATTTCTTTACATAGATTT 622
|||||
Db 195518 AATATTTCTTTACATAGATTT 195495

RESULT 13
AR290033/c 47 bp DNA linear PAT 12-JUN-2003
LOCUS AR290033
DEFINITION Sequence 1768 from patent US 6537751.
ACCESSION AR290033
VERSION AR290033.1 GI:31677317
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 47)
Cohen,D., Chumakov,I. and Blumenfeld,M.
TITLE Biallelic markers for use in constructing a high density
disequilibrium map of the human genome
JOURNAL Patent: US 6537751-A 1768 25-MAR-2003;
FEATURES
source
1..47
/organism="unknown"

ORIGIN
/mol_type="genomic DNA"

Query Match 2.1%; Score 23; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 531 AATGACATTTTACACCTTTCAA 553
|||||
Db 47 AATGACATTTTACACCTTTCAA 25

RESULT 14
AX114392 47 bp DNA linear PAT 11-MAY-2001
LOCUS AX114392
DEFINITION Sequence 61 from Patent WO0129257.
ACCESSION AX114392
VERSION AX114392.1 GI:14031356
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Schork,N. and Skierczynski,B.
TITLE Methods of genetic cluster analysis and use thereof
JOURNAL Patent: WO 0129257-A 61 26-APR-2001;
GENSET (PR)

FEATURES
source
1..47
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
24
/note="99-622-95 : polymorphic base A or G"

ORIGIN

Query Match 2.1%; Score 23; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 531 AATGACATTTTACACCTTTCAA 553
|||||
Db 47 AATGACATTTTACACCTTTCAA 25

RESULT 15
AX281468/c 6070 bp DNA linear PAT 02-NOV-2001
LOCUS AX281468
DEFINITION Sequence 132 from Patent WO0177376.
ACCESSION AX281468
VERSION AX281468.1 GI:16608723
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with metastasis
JOURNAL Patent: WO 0177376-A 132 18-OCT-2001;
Epigenomics AG (DE)
FEATURES
source
1..6070
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match 2.1%; Score 23; DB 6; Length 6070;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 480 TTTATTATTATTATTATAA 502
| | | | | | | | | | | | | | | |
Db 2817 TTTATTATTATTATTATAA 2795

Search completed: May 26, 2004, 10:29:41
Job time : 4181.21 secs

THIS PAGE LEFT BLANK

Db 1021 ATATCTGTGGGATTTGTTCTTCAAGTATGATTAACATGTAATATCATATTAATCTGTG 1080
 Qy 1081 ATGGGAAAAACAAGGACCCCTAATTTGAAACACCTGTGTAGCGAGAGGAGGAGG 1140
 Db 1081 ATGGGAAAAACAAGGACCCCTAATTTGAAACACCTGTGTAGCGAGAGGAGGAGG 1140
 Qy 1141 GGGCTGTGCGCACTCAGAGCGGAGGCTGAGAGGCGGCGTCCCTTGCAAGAGACTGCA 1200
 Db 1141 GGGCTGTGCGCACTCAGAGCGGAGGCTGAGAGGCGGCGTCCCTTGCAAGAGACTGCA 1200
 Qy 1201 GTGAGCAGATGGGAGCACTGAGCTGCCCGGAGACTGTGGGCTGAGCTGCTTAACCTGG 1260
 Db 1201 GTGAGCAGATGGGAGCACTGAGCTGCCCGGAGACTGTGGGCTGAGCTGCTTAACCTGG 1260
 Qy 1261 GCGGAGGCTCTGCAAGAAATTAAGACCTCCGATTAACCTTAACCACTTTCTCACTGCTC 1320
 Db 1261 GCGGAGGCTCTGCAAGAAATTAAGACCTCCGATTAACCTTAACCACTTTCTCACTGCTC 1320
 Qy 1321 TAATTTGTGTGATCCCGGCGCCGAGAGGCTTGTGAGCAGAGTGTGCGGCTTCCAGGAGC 1380
 Db 1321 TAATTTGTGTGATCCCGGCGCCGAGAGGCTTGTGAGCAGAGTGTGCGGCTTCCAGGAGC 1380
 Qy 1381 TCCAGCAGCCTTTAAACCTGAACCGGCGCACGTCGCGCCGAGAGGAGCAAAAGAGGC 1440
 Db 1381 TCCAGCAGCCTTTAAACCTGAACCGGCGCACGTCGCGCCGAGAGGAGCAAAAGAGGC 1440
 Qy 1441 ACCGAGACCTCTCCCGGCGGAGCAACCACTTCAACCCAGTTCCTGCTAGTCCGCAACCT 1500
 Db 1441 ACCGAGACCTCTCCCGGCGGAGCAACCACTTCAACCCAGTTCCTGCTAGTCCGCAACCT 1500
 Qy 1501 CCTTCCCGGCTCCGAGCGGAGCGGAGCGGAGGAGCATGCGCAGTGTGCGGAGCGGAGCT 1560
 Db 1501 CCTTCCCGGCTCCGAGCGGAGCGGAGCGGAGGAGCATGCGCAGTGTGCGGAGCGGAGCT 1560
 Qy 1561 GCGGCGCGCAGCAGAGGAGTGTGTAATGCAACTGTGCGGCCCAACCAACCAATCAAGA 1620
 Db 1561 GCGGCGCGCAGCAGAGGAGTGTGTAATGCAACTGTGCGGCCCAACCAACCAATCAAGA 1620
 Qy 1621 CGTGTATTTGCGCGGAGGTGAACCTATGCAACGCGGAGCAACCAATCAAGAGCGGTTGT 1680
 Db 1621 CGTGTATTTGCGCGGAGGTGAACCTATGCAACGCGGAGCAACCAATCAAGAGCGGTTGT 1680
 Qy 1681 TGCGGCGGAGCGGCGGAGCGGAGCGGAGGAGTGTGCGGAGTGTGCGGCGGAGG 1740
 Db 1681 TGCGGCGGAGCGGCGGAGCGGAGCGGAGGAGTGTGCGGAGTGTGCGGCGGAGG 1740
 Qy 1741 GAGCATCCCTGAGCCATGATCCGAGAGGCGCGGAGTGTGCGGCGGAGG 1800
 Db 1741 GAGCATCCCTGAGCCATGATCCGAGAGGCGCGGAGTGTGCGGCGGAGG 1800
 Qy 1801 CGGCGCAGCAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1860
 Db 1801 CGGCGCAGCAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1860
 Qy 1861 TGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1920
 Db 1861 TGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1920
 Qy 1921 CGGCGCTGTGCGGAGCTGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1980
 Db 1921 CGGCGCTGTGCGGAGCTGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1980
 Qy 1981 GCGGCGGAGCTGAGGAGAGTGTGCTGTCTCTGCTGTCTCAACGTAATGCGCTTACC 2040
 Db 1981 GCGGCGGAGCTGAGGAGAGTGTGCTGTCTCTGCTGTCTCAACGTAATGCGCTTACC 2040
 Qy 2041 TGCTGCCAGAGGTGTCTCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 2100
 Db 2041 TGCTGCCAGAGGTGTCTCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 2100
 Qy 2101 GCGGCGGAGCTCTCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 2160
 Db 2101 GCGGCGGAGCTCTCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 2160

Db 2101 GCGGCGGAGCTCTCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCT 2160
 Qy 2161 ACTGCTGTACCAAGAGATGATGCTCTTCTTCTTCAAGAAATTAACCGGAGTCCAGGTGA 2220
 Db 2161 ACTGCTGTACCAAGAGATGATGCTCTTCTTCTTCAAGAAATTAACCGGAGTCCAGGTGA 2220
 Qy 2221 GCGGCTTCCGCTCCCGGAGTGTGCGGCTTCCAGGAGTCTCCGGGAGCGGAGCTTCTCC 2280
 Db 2221 GCGGCTTCCGCTCCCGGAGTGTGCGGCTTCCAGGAGTCTCCGGGAGCGGAGCTTCTCC 2280
 Qy 2281 GCTTCCCCCAAGCTGTGAGAGGTTCACCGGAGCGGAGCGGAGCGGAGCGGAGCT 2340
 Db 2281 GCTTCCCCCAAGCTGTGAGAGGTTCACCGGAGCGGAGCGGAGCGGAGCGGAGCT 2340
 RESULT 2
 BD196564
 LOCUS BD196564 56520 bp DNA linear PAT 17-JUN-2003
 DEFINITION Prostatic cancer gene.
 ACCESSION BD196564 GI:33006334
 VERSION BD196564.1
 KEYWORDS JP 2002516657-A/153.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueleret, L.
 AUTHORS Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueleret, L.
 TITLE Prostatic cancer gene
 JOURNAL Patent: JP 2002516657-A 153 11-JUN-2002;
 GENSER
 COMMENT OS Homo sapiens (human)
 PN JP 2002516657-A/153
 PD 11-JUN-2002
 PF 22-DEC-1998 JP 2000525562
 PR 22-DEC-1997 US 08/996306, 09-SEP-1998 US 60/099658 PT
 DANIEL COHEN, MARTA BLUMENFELD, ILYA CHUMAKOV, LYDIE BOUGUELERET PC
 C12N15/09, C12N15/09, A01K67/027, C07K14/47, C07K16/18, C12N1/15, PC
 C12N1/19
 PC C12N1/21, C12N5/10, C12N5/10, C12P21/08, C12Q1/68, G01N33/50 PC
 PC C12N15/00, C12N5/00
 PC C12N5/00, C12N15/00
 CC exon1
 CC exon2
 CC exon3
 CC exon4
 CC exon5
 CC exon6
 CC exon7
 CC exon8
 CC AATTA
 CC upstream amplification primer 5-63
 CC downstream amplification primer 5-63 , complement CC
 CC downstream amplification primer 99-622
 CC upstream amplification primer 99-622 , complement CC
 CC downstream amplification primer 99-621
 CC upstream amplification primer 99-621 , complement CC
 CC downstream amplification primer 99-619
 CC upstream amplification primer 99-619 , complement CC
 CC downstream amplification primer 4-76
 CC upstream amplification primer 4-76 , complement CC
 CC downstream amplification primer 4-77
 CC upstream amplification primer 4-77 , complement CC
 CC downstream amplification primer 4-71
 CC upstream amplification primer 4-71 , complement CC
 CC downstream amplification primer 4-72
 CC upstream amplification primer 4-72 , complement CC
 CC downstream amplification primer 4-73
 CC upstream amplification primer 4-73 , complement CC
 CC downstream amplification primer 4-73
 CC upstream amplification primer 4-73 , complement CC
 CC downstream amplification primer 99-610
 CC upstream amplification primer 99-610 , complement CC
 CC downstream amplification primer 99-609
 CC upstream amplification primer 99-609 , complement CC

upstream amplification primer 4-90	CC	downstream amplification primer 4-90	complement CC	FT	primer_bind	2505	.2525
CC downstream amplification primer 4-90	complement CC	FT	primer_bind	4091	.4111		
CC upstream amplification primer 99-607	CC	upstream amplification primer 99-607	complement CC	FT	primer_bind	4528	.4546
CC downstream amplification primer 99-607	complement CC	FT	primer_bind	5475	.5495		
CC upstream amplification primer 99-602	CC	upstream amplification primer 99-602	complement CC	FT	primer_bind	5927	.5947
CC downstream amplification primer 99-602	complement CC	FT	primer_bind	8127	.8144		
CC upstream amplification primer 99-600	CC	upstream amplification primer 99-600	complement CC	FT	primer_bind	8560	.8578
CC downstream amplification primer 99-600	complement CC	FT	primer_bind	11622	.11639		
CC upstream amplification primer 99-598	CC	upstream amplification primer 99-598	complement CC	FT	primer_bind	12018	.12037
CC downstream amplification primer 99-598	complement CC	FT	primer_bind	11930	.11947		
CC upstream amplification primer 99-592	CC	upstream amplification primer 99-592	complement CC	FT	primer_bind	12339	.12358
CC downstream amplification primer 99-592	complement CC	FT	primer_bind	12915	.12932		
CC upstream amplification primer 99-217	CC	upstream amplification primer 99-217	complement CC	FT	primer_bind	13317	.13334
CC downstream amplification primer 99-217	complement CC	FT	primer_bind	13216	.13233		
CC upstream amplification primer 5-47	CC	upstream amplification primer 5-47	complement CC	FT	primer_bind	13617	.13636
CC downstream amplification primer 5-47	complement CC	FT	primer_bind	13547	.13564		
CC upstream amplification primer 99-589	CC	upstream amplification primer 99-589	complement CC	FT	primer_bind	13962	.13981
CC downstream amplification primer 99-589	complement CC	FT	primer_bind	15994	.16011		
CC upstream amplification primer 99-588	CC	upstream amplification primer 99-588	complement CC	FT	primer_bind	16463	.16480
CC downstream amplification primer 99-588	complement CC	FT	primer_bind	17304	.17324		
CC upstream amplification primer 99-12899	CC	upstream amplification primer 99-12899	complement CC	FT	primer_bind	17814	.17832
CC downstream amplification primer 99-12899	complement CC	FT	primer_bind	18008	.18027		
CC upstream amplification primer 4-12	CC	upstream amplification primer 4-12	complement CC	FT	primer_bind	18423	.18442
CC downstream amplification primer 4-12	complement CC	FT	primer_bind	18699	.18716		
CC upstream amplification primer 99-582	CC	upstream amplification primer 99-582	complement CC	FT	primer_bind	19164	.19182
CC downstream amplification primer 99-582	complement CC	FT	primer_bind	22589	.22609		
CC upstream amplification primer 99-576	CC	upstream amplification primer 99-576	complement CC	FT	primer_bind	23111	.23129
CC downstream amplification primer 99-576	complement CC	FT	primer_bind	25098	.25118		
CC upstream amplification primer 4-13	CC	upstream amplification primer 4-13	complement CC	FT	primer_bind	25657	.25674
CC downstream amplification primer 4-13	complement CC	FT	primer_bind	26537	.26557		
CC upstream amplification primer 99-12903	CC	upstream amplification primer 99-12903	complement CC	FT	primer_bind	27022	.27040
CC downstream amplification primer 99-12903	complement CC	FT	primer_bind	32262	.32281		
CC upstream amplification primer 5-56	CC	upstream amplification primer 5-56	complement CC	FT	primer_bind	32823	.32841
CC downstream amplification primer 5-56	complement CC	FT	primer_bind	34215	.34233		
CC upstream amplification primer 4-61	CC	upstream amplification primer 4-61	complement CC	FT	primer_bind	34624	.34644
CC downstream amplification primer 4-61	complement CC	FT	primer_bind	34473	.34491		
CC upstream amplification primer 4-62	CC	upstream amplification primer 4-62	complement CC	FT	primer_bind	34916	.34936
CC downstream amplification primer 4-62	complement CC	FT	primer_bind	34702	.34722		
CC upstream amplification primer 4-63	CC	upstream amplification primer 4-63	complement CC	FT	primer_bind	35182	.35200
CC downstream amplification primer 4-63	complement CC	FT	primer_bind	39591	.39611		
CC upstream amplification primer 4-64	CC	upstream amplification primer 4-64	complement CC	FT	primer_bind	39971	.39991
CC downstream amplification primer 4-64	complement CC	FT	primer_bind	40531	.40549		
CC upstream amplification primer 4-65	CC	upstream amplification primer 4-65	complement CC	FT	primer_bind	40932	.40950
CC downstream amplification primer 4-65	complement CC	FT	primer_bind	40629	.40649		
CC upstream amplification primer 4-67	CC	upstream amplification primer 4-67	complement CC	FT	primer_bind	41058	.41078
CC downstream amplification primer 4-67	complement CC	FT	primer_bind	45729	.45746		
CC upstream amplification primer 5-50	CC	upstream amplification primer 5-50	complement CC	FT	primer_bind	46186	.46203
CC downstream amplification primer 5-50	complement CC	FT	primer_bind	47879	.47896		
CC upstream amplification primer 5-71	CC	upstream amplification primer 5-71	complement CC	FT	primer_bind	48217	.48236
CC downstream amplification primer 5-71	complement CC	FT	primer_bind	48902	.48922		
CC upstream amplification primer 5-30	CC	upstream amplification primer 5-30	complement CC	FT	primer_bind	49331	.49351
CC downstream amplification primer 5-30	complement CC	FT	primer_bind	49830	.49848		
CC upstream amplification primer 5-58	CC	upstream amplification primer 5-58	complement CC	FT	primer_bind	50271	.50290
CC downstream amplification primer 5-58	complement CC	FT	primer_bind	50172	.50189		
CC upstream amplification primer 5-53	CC	upstream amplification primer 5-53	complement CC	FT	primer_bind	50573	.50591
CC downstream amplification primer 5-53	complement CC	FT	primer_bind	50541	.50560		
CC upstream amplification primer 5-60	CC	upstream amplification primer 5-60	complement CC	FT	primer_bind	50940	.50959
CC downstream amplification primer 5-60	complement CC	FT	primer_bind	50555	.50572		
CC upstream amplification primer 5-68	CC	upstream amplification primer 5-68	complement CC	FT	primer_bind	50964	.50983
CC downstream amplification primer 5-68	complement CC	FT	primer_bind	50774	.50792		
CC upstream amplification primer 5-66	CC	upstream amplification primer 5-66	complement CC	FT	primer_bind	51183	.51202
CC downstream amplification primer 5-66	complement CC	FT	primer_bind	51146	.51165		
CC upstream amplification primer 5-62	CC	upstream amplification primer 5-62	complement CC	FT	primer_bind	51479	.51496
CC downstream amplification primer 5-62	complement CC	FT	primer_bind	51593	.51610		
Location/Qualifiers				FT	primer_bind	29734	.29744
FT exon	2001	.2216		FT	primer_bind	51167	.51185
FT exon	18196	.18265		FT	primer_bind	51677	.51687
FT exon	23716	.23831		FT	primer_bind	51387	.51403
FT exon	25570	.25659		FT	primer_bind	51826	.51843
FT exon	34668	.34758		FT	primer_bind	51772	.51789
FT exon	40685	.40843		FT	primer_bind	52199	.52217
FT exon	48067	.48190					
FT exon	50179	.54519					
FT polyA_signal	54493	.54498					
FT primer_bind	1991	.2008					

Query Match

100.0%; Score 2324; DB 6; Length 56520;

QY	2161	ACTGCGCTCTACCAAGAGATGCTGCTCTTCTCTTCGGAATTACACCGGGGTCCAGGTGA	2220
Db	2161	ACTGCGCTCTACCAAGAGATGCTGCTCTTCTCTTCGGAATTACACCGGGGTCCAGGTGA	2220
QY	2221	GCCGCTCCCGCTCCGGGCTCCGGCGTCCACCCGAGTCCCGGGGGCGCGGACTCTCC	2280
Db	2221	GCGGCTCTCCGCTCCGGGCTCCGGCGTCCAGCCGAGTCCCGGGGGCGCGGACTCTCC	2280
QY	2281	GCTTCCCCCACAAGCTGGCGAGGGTCACTCCGGGCGGCGCGGCGGAC	2324
Db	2281	GCTTCCCCCACAAGCTGGCGAGGGTCACTCCGGGCGGCGGCGGCGGAC	2324
RESULT 3			
BD196565			
LOCUS	BD196565	2000 bp	DNA linear PAT 17-JUL-2003
DEFINITION	Prostatic cancer gene.		
ACCESSION	BD196565		
VERSION	BD196565.1 GI:33006335		
KEYWORDS	JP 2002516657-A/154.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueleret, L.		
TITLE	Prostatic cancer gene		
JOURNAL	Patent: JP 2002516657-A 154 11-JUN-2002;		
COMMENT			
OS	Homo sapiens (human)		
PN	JP 2002516657-A/154		
PD	11-JUN-2002		
PF	22-DEC-1998 JP 2000525562		
PR	22-DEC-1997 US 08/996306, 09-SEP-1998 US 60/099658 PI		
DANIEL COHEN, MARTA BLUMENFELD, ILYA CHUMAKOV, IYDIE BOUGUELERET	PC		
CI2N15/09, CI2N15/09, A01K67/027, C07K14/47, C07K16/18, CI2N1/15, PC			
CI2N1/19, PC			
CI2N1/21, CI2N5/10, CI2N5/10, CI2P21/08, CI2Q1/68, G01N33/50 PC			
CI2N15/00, CI2N5/00, PC			
CI2N5/00, CI2N15/00			
CC	Prostatic cancer gene		
FT	Key		
FT	Location/Qualifiers		
FT	1. 2000		
FT	location/Qualifiers		
FT	1. 2000		
FT	/organism="Homo sapiens (human)".		
FEATURES			
source	1. 2000		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	86.1%;	Score 2000;	DB 6; Length 2000;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2000; Conservative	0;	Mismatches	0; Indels 0; Gaps 0
QY	1	GTGATCTGTGACTGTTGCGAGGAAGAGAGAGCGGAGCAGAGACAGACATTAATCTGATA	60
Db	1	GTGATCTGTGACTGTTGCGAGGAAGAGAGAGCGGAGCAGAGACAGACATTAATCTGATA	60
QY	61	GTGAGAGAGTGGGTTTGGAGTAAAGAGGGAACAGAGAAAGTTAAGTTCTGTGTTTCA	120
Db	61	GTGAGAGAGTGGGTTTGGAGTAAAGAGGGAACAGAGAAAGTTAAGTTCTGTGTTTCA	120
QY	121	TGGCAAAACATTCACAAAAGTTTACAACTTCGTGACTTAACAGTAATCTGGGGTGATTCAC	180
Db	121	TGGCAAAACATTCACAAAAGTTTACAACTTCGTGACTTAACAGTAATCTGGGGTGATTCAC	180
QY	181	AACAAATTTACACATTAACACATATTTACTGACTTTATACACGCAATCTTAACGTGAAC	240
Db	181	AACAAATTTACACATTAACACATATTTACTGACTTTATACACGCAATCTTAACGTGAAC	240
QY	241	ACAGAACCTCGCTTATCTTTTGGCAACAGCTTCTAGTGTAGAGATGTCTGCTCAGTTA	300

Db	241	ACGAACTCTTATCTTTTGGCAACTGTTCTTAGTGTGAGATGTCTGTCTAGTTA	300
Qy	301	AAGAAACATTAAGAGACATTAGTTGTGACACACTGTCCACACCCGTGACTTTTTCACCA	360
Db	301	AAGAAACATTAAGAGACATTAGTTGTGACACTGTCCACACCCGTGACTTTTTCACCA	360
Qy	361	GTACTAAACCTTAGTGTCTTCAAGTACAGGGGGAATGACAGCCACAGAAAGAGAAGCT	420
Db	361	GTACTAAACCTTAGTGTCTTCAAGTACAGGGGGAATGACAGCCACAGAAAGAGAAGCT	420
Qy	421	CCCTTTACGTGTGAATGCTTCTGTGTGGCCCTTCAAACTCTGTACTGTAGAGATCTCCA	480
Db	421	CCCTTTACGTGTGAATGCTTCTGTGTGGCCCTTCAAACTCTGTACTGTAGAGATCTCCA	480
Qy	481	TTTCACTGTGCTTTGTCCCCAAGGTCAATCATACCAATGATGTGTATTTTGAATGTAA	540
Db	481	TTTCACTGTGCTTTGTCCCCAAGGTCAATCATACCAATGATGTGTATTTTGAATGTAA	540
Qy	541	TCATGTATTAAGAAAGTAGCTACCATCTGTGGCCCTGATTGAACCTTCCACTGAATACC	600
Db	541	TCATGTATTAAGAAAGTAGCTACCATCTGTGGCCCTGATTGAACCTTCCACTGAATACC	600
Qy	601	GTCTCTCCCTTAAGGTAGCAAGGTTTCATTATGTGTGTGTGTGGGAGGGGGGGGAATA	660
Db	601	GTCTCTCCCTTAAGGTAGCAAGGTTTCATTATGTGTGTGTGTGGGAGGGGGGGGAATA	660
Qy	661	TAT	720
Db	661	TAT	720
Qy	721	ACTATCCCTTGAAGAGGTTACATATTAACCATTTTTCACACAGCCAAAGGGGAGAGAA	780
Db	721	ACTATCCCTTGAAGAGGTTACATATTAACCATTTTTCACACAGCCAAAGGGGAGAGAA	780
Qy	781	AGATCCAAAGGCTCGTGTGAGTGTCTTAAATCATCAATTAACAAGTTATCCACCTTGTA	840
Db	781	AGATCCAAAGGCTCGTGTGAGTGTCTTAAATCATCAATTAACAAGTTATCCACCTTGTA	840
Qy	841	GCTTTTAGTGAAGCTACAAAAGTATGCTTTTATGTGATTTACATGTGTGACCGCACTAC	900
Db	841	GCTTTTAGTGAAGCTACAAAAGTATGCTTTTATGTGATTTACATGTGTGACCGCACTAC	900
Qy	901	TTTATATCTACAGAAAAAAACGAGGCTCCTTAATAAAAAAAATCGAAAAACAAGTCCA	960
Db	901	TTTATATCTACAGAAAAAAACGAGGCTCCTTAATAAAAAAAATCGAAAAACAAGTCCA	960
Qy	961	CAGACTCTGAGGAATGAGCAAGCAAGTGAATTCGAAAAAGGCTTAATTAACAGTATGAA	1020
Db	961	CAGACTCTGAGGAATGAGCAAGCAAGTGAATTCGAAAAAGGCTTAATTAACAGTATGAA	1020
Qy	1021	ATATCTCTGTGGAGATTGTTCTTCAGCTATGTATGAATGAATTAATCAATCATTACTGTG	1080
Db	1021	ATATCTCTGTGGAGATTGTTCTTCAGCTATGTATGAATGAATTAATCAATCATTACTGTG	1080
Qy	1081	ATGGGGAATAACAAGCACTTAATTCGAAACACCTGTGTAGCGAGAGACGGGACGAGAGG	1140
Db	1081	ATGGGGAATAACAAGCACTTAATTCGAAACACCTGTGTAGCGAGAGACGGGACGAGAGG	1140
Qy	1141	GGCTGCTCGCACTCAGACGCGAGGCTGAGGAGCGCGCTCCCTTCGAAAGGACTGGCA	1200
Db	1141	GGCTGCTCGCACTCAGACGCGAGGCTGAGGAGCGCGCTCCCTTCGAAAGGACTGGCA	1200
Qy	1201	GTGACACATATGGGACACTCGAGCTGCGCCCGCAGACTGTGGCCGAGCTGTCAACACTGG	1260
Db	1201	GTGACACATATGGGACACTCGAGCTGCGCCCGCAGACTGTGGCCGAGCTGTCAACACTGG	1260
Qy	1261	GCCCAAGTGCCTGCAAGATTAAGACTTCGATTAAGTTAACCCCACTTCTCACTGCTC	1320
Db	1261	GCCCAAGTGCCTGCAAGATTAAGACTTCGATTAAGTTAACCCCACTTCTCACTGCTC	1320
Qy	1321	TATTTGTGTGATCCCCGGCCCAAGGGCTTGTGAGCAGAGGTGTGGCTTCCAGGACGC	1380
Db	1321	TATTTGTGTGATCCCCGGCCCAAGGGCTTGTGAGCAGAGGTGTGGCTTCCAGGACGC	1380


```

Db 1321 TAATTGTGATCCCGCGCCAGAGGGCTTGTGACAGAGTGCAGCTTCCAGGAGC 1380
Qy 1381 TCACAGCAGCCCTTAACCTGACCGCGCGAGCTCCGCGCGAGGAGGAGCAAGAGGC 1440
Db 1381 TCACAGCAGCCCTTAACCTGACCGCGCGAGCTCCGCGCGAGGAGGAGCAAGAGGC 1440
Qy 1441 ACCCGGAGCCCTCTCCCGGAGCAGCAGCAGCTTCAACCCAGTTCCGTCAGTCGCGCACACT 1500
Db 1441 ACCCGGAGCCCTCTCCCGGAGCAGCAGCAGCTTCAACCCAGTTCCGTCAGTCGCGCACACT 1500
Qy 1501 CCCTTCCCGCGCTCCGAGCCGCGAGCTGGAGGAGCATGCGCAGTGGCCGAGCGGGTT 1560
Db 1501 CCCTTCCCGCGCTCCGAGCCGCGAGCTGGAGGAGCATGCGCAGTGGCCGAGCGGGTT 1560
Qy 1561 GCGCGCGCAGCAGAGTGTGTCTGCACTGTGCGCCCAACCAACCAATCAAGAGA 1620
Db 1561 GCGCGCGCAGCAGAGTGTGTCTGCACTGTGCGCCCAACCAACCAATCAAGAGA 1620
Qy 1621 CGTGTATTGCGCGCGAGGTGAGTATGCAACCGCGCGAGCAGCAGTGT 1680
Db 1621 CGTGTATTGCGCGCGAGGTGAGTATGCAACCGCGCGAGCAGCAGTGT 1680
Qy 1681 TGCCGCGAGCCCTCTCCCGGAGCAGGAGGATGTGGCAGTGGGTGATGGGGTGT 1740
Db 1681 TGCCGCGAGCCCTCTCCCGGAGCAGGAGGATGTGGCAGTGGGTGATGGGGTGT 1740
Qy 1741 GAGCATCCCTGAGCATGATCCGCGAGGCGCGGGTTCCTTGTGCGCGCGAG 1800
Db 1741 GAGCATCCCTGAGCATGATCCGCGAGGCGCGGGTTCCTTGTGCGCGCGAG 1800
Qy 1801 CGCGCGCAGCAGCAGCCCGGAGCTGCTACCCGCGCGCGCGCGCGCGCGCATGCGGC 1860
Db 1801 CGCGCGCAGCAGCAGCCCGGAGCTGCTACCCGCGCGCGCGCGCGCGCGCATGCGGC 1860
Qy 1861 TGGGGGCGAGAGCTGGAGAGCGGGTGGCGGCGCGCGCGCGCGCGCGCGGTGATGG 1920
Db 1861 TGGGGGCGAGAGCTGGAGAGCGGGTGGCGGCGCGCGCGCGCGCGCGCGGTGATGG 1920
Qy 1921 CGCGCTGTGGCGCGGAGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGAGTCCCT 1980
Db 1921 CGCGCTGTGGCGCGGAGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGAGTCCCT 1980
Qy 1981 GCGCGCGAGCTGAGAGAGTGG 2000
Db 1981 GCGCGCGAGCTGAGAGAGTGG 2000

RESULT 4
AC100813/c 165799 bp DNA linear HTG 27-MAR-2003
LOCUS Homo sapiens chromosome 8 clone CTD-2541M15 map 8, 4 unordered
DEFINITION
pieces.
AC100813
AC100813.2 GI:22381986
VERSION HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 165799)
Britten,B., Linton,L., Nusbbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhaltier,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kelle,C., LaRocque,K.,

```

```

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margulis,N., Mathew,C.,
McCarthy,M., Mewan,P., McKernan,K., McSheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnuppack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmmer,A. and Zody,M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165799)
Britten,B., Nusbbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhaltier,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schnuppack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmmer,A. and Zody,M.
Direct Submission
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:117048183.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21488
Center clone name: 2541_M_15
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 62727: contig of 62727 bp in length
* 62827: gap of 100 bp
* 62828 73470: contig of 10643 bp in length
* 73471 73570: gap of 100 bp
* 73571 81515: contig of 7945 bp in length
* 81516 165799: gap of 100 bp
* 81616 165799: contig of 84184 bp in length.
Location/Qualifiers
1. 165799
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"

```

```

/clone="CTD-254IM15"
/clone_lib="CITD2 Human BAC"
ORIGIN

```

Query Match	67.3%	Score 1565	DB 2	Length 165799
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1665	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

[illegible]

Db	54357	AGACGTGTTATTTCGCCCCGACGGTGGAACTATGGCAACGGGCGACCAATCAGAAAGCGCGT	54298
QY	1678	TGTTGCCGCGAGACCCCTTGCCCTCCGCGACGGGGATGTGGCGATGGGTGAGGGTCA.TGGGG	1737
Db	54297	TGTTGCCGCGAGACCCCTTGCCCTCCGCGACGGGGGATGTGGCGATGGGTGAGGGTCA.TGGGG	54238
QY	1738	TGTGAGCATATCCCGAGCCCATTCGATCCGGGAGGGGCGCGGGTTCCCTGTGGTTTGGCGCGG	1797
Db	54237	TGTGAGCATATCCCTGAGCATTCGATCCGGGAGGGGCGCGGGTTCCCTGTGGTTTGGCGCGG	54178
QY	1798	GAGCGCGCACGCGACGCCCGCACTCGCTCAACCGAGCCCGCGCGCGCGCGCGCCCATGC	1857
Db	54177	GAGCGCGCACGCGACGCCCGCACTCGCTCAACCGAGCCCGCGCGCGCGCGCGCCCATGC	54118
QY	1858	GCGTGGGGGCGGAGGCTGGAGCGGGTGGCGGGCGGGCGCGCGCGCGCGCGCGCGTGTAT	1917
Db	54117	GCGTGGGGGCGGAGGCTGGAGCGGGTGGCGGGCGGGCGCGCGCGCGCGCGCGCGTGTAT	54058
QY	1918	TGGCGCGCTGTGGCGCGCGCATGAGGCGCGGGAGAGGGCGGGGGAGCGACGCGGAGAGTC	1977
Db	54057	TGGCGCGCTGTGGCGCGCGCATGAGGCGCGGGAGAGGGCGGGGGAGCGACGCGGAGAGTC	53998
QY	1978	GCTGCAGCCAGCTGAGAAAGATGCTCTTCCCTGTGTCTCCACACAGTACTCATAGCGCT	2037
Db	53997	GCTGCAGCCAGCTGAGAAAGATGCTCTTCCCTGTGTCTCCACACAGTACTCATAGCGCT	53938
QY	2038	ACCTGCTGCCAGCGCTGTGCTCTTGGGCACGGCGCCACCTTACGTGTGGCTTGGGGGG	2097
Db	53937	ACCTGCTGCCAGCGCTGTGCTCTTGGGCACGGCGCCACCTTACGTGTGTGGCTTGGGGGG	53878
QY	2098	TCTGGGCGGCGCTCTCGGCTTCCTGTCGCGCGCGCGCGCTTCAACAAGGGCTGAGACGACGGC	2157
Db	53877	TCTGGGCGGCGCTCTCGGCTTCCTGTCGCGCGCGCGCGCTTCAACAAGGGCTGAGACGACGGC	53818
QY	2158	TGTACTGCGCTACACGAGCATGTGCTCTTCTTCTTCGAGATTATCAACCGGGGTCCAGG	2217
Db	53817	TGTACTGCGCTACACGAGCATGTGCTCTTCTTCTTCGAGATTATCAACCGGGGTCCAGG	53758
QY	2218	TGAGCGGCGCTCCGCGTCCGGGGTCTTGCGGCTCCACCGAGCTTCCGGGGGGCGGACCTTC	2277
Db	53757	TGAGCGGCGCTCCGCGTCCGGGGTCTTGCGGCTCCACCGAGCTTCCGGGGGGCGGACCTTC	53698
QY	2278	TTCGCTCCCCACAGCTGGCGAGGGGTCAACCGCGCGGCGCGCGCGGAC	2324
Db	53697	TTCGCTCCCCACAGCTGGCGAGGGGTCAACCGCGCGGCGCGCGCGGCGGAC	53651

RESULT	5
LOCUS	AF287957
DEFINITION	168136 bp DNA linear PRI 08-JUL-2003
ACCESSION	AF287957
VERSION	AF287957
KEYWORDS	AF287957.6 GI:32469520
SOURCE	HMG.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 168136) Reichwald,K., Baunzgart,C., Blechschmidt,K., Dette,M., Jahn,N., Meinzel,U., Polley,A., Schlubel,M.B., Taudien,S., Wen,G. and Rosenthal,A. Chromosome 8 genomic sequence Unpublished 2 (bases 1 to 168136) Genome Sequencing Center Jena. Direct Submission Submitted (14-JUL-2000) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany 3 (bases 1 to 168136) Reichwald,K. and Platzer,M.
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	

TITLE Direct Submission
JOURNAL Submitted (23-AUG-2001) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE 4 (bases 1 to 168136)
AUTHORS Lagemann,D. and Platzer,M.
JOURNAL Direct Submission
TITLE Submitted (29-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE 5 (bases 1 to 168136)
AUTHORS Lagemann,D. and Platzer,M.
JOURNAL Direct Submission
TITLE Submitted (03-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE 6 (bases 1 to 168136)
AUTHORS Lagemann,D. and Platzer,M.
JOURNAL Direct Submission
TITLE Submitted (09-JUL-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 7 (bases 1 to 168136)
AUTHORS Lagemann,D. and Platzer,M.
JOURNAL Direct Submission
TITLE Submitted (08-JUL-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Jul 8, 2003 this sequence version replaced gi:21717104.
COMMENT Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: <http://genome.imb-jena.de/>
Contact: gascj-submit@genome.imb-jena.de
Project Information
Center project name: H351
Center clone name: CTD-2541M15
Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 167707 bases at least Q40
Consensus quality: 168123 bases at least Q30
Consensus quality: 168136 bases at least Q20
Quality coverage: 6.18x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phrap quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one pUC18 subclone; and the
assembly was confirmed by restriction digest.

Neighboring sequence information:
This clone is overlapped by GSI-24F4, GSI-24F4.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
SOURCE
1..168136
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="p22-p21"
/clone="CTD-2541M15"
165266..168136
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="taxon:9606"
/clone="GSI-24F4"
/note="overlapping clone, overlapping clone"
misc_feature
6418
/note="low quality region"
7442..7645
/note="single stranded/single chemistry region"
28243..28507
/note="single stranded/single chemistry region"
28499
/note="low quality region"
28501
/note="low quality region"
34532..34843
/note="single stranded/single chemistry region"
35987..36115
/note="single stranded/single chemistry region"
37530
/note="low quality region"
38752..38820
/note="single stranded/single chemistry region"
38995..39017
/note="single stranded/single chemistry region"
44702..44825
/note="single stranded/single chemistry region"
45312..45406
/note="single stranded/single chemistry region"
47601..47995
/note="single stranded/single chemistry region"
59027..59136
/note="single stranded/single chemistry region"
61319
/note="low quality region"
62590..62726
/note="single stranded/single chemistry region"
66851..66899
/note="single stranded/single chemistry region"
69911..70188
/note="single stranded/single chemistry region"
75977..76021
/note="single stranded/single chemistry region"
78279..78386
/note="single stranded/single chemistry region"
78675
/note="low quality region"
81011..81279
/note="single stranded/single chemistry region"
81070..81071
/note="low quality region"
81743..82001
/note="single stranded/single chemistry region"
81872
/note="low quality region"
84434..85028
/note="pcr product sequence only"
84932..85028
/note="single stranded/single chemistry region"
86073..86097
/note="single stranded/single chemistry region"
86609
/note="low quality region"
89636..89783
/note="single stranded/single chemistry region"
90439..90940
/note="single stranded/single chemistry region"
95593..95693
/note="single stranded/single chemistry region"
98827..98848
/note="single stranded/single chemistry region"
98860
/note="low quality region"
99495..99715
/note="single stranded/single chemistry region"

QY 914 GAAAAAACGAGGCTCTTATTTAAAAAATCAGAAAAGCTCCAGACAGCTCGAGGA 973
DB 73173 GAAAAAACGAGGCTCTTATTTAAAAAATCAGAAAAGCTCCAGACAGCTCGAGGA 73232
QY 974 AATGAAGCAAGAGTAAATCTGAAAAGGCTATATTAACAGTATGAAAAATCTCTGTGGG 1033
DB 73232 AATGAAGCAAGAGTAAATCTGAAAAGGCTATATTAACAGTATGAAAAATCTCTGTGGG 73232
QY 1034 ATTGTTCTTACGCTATGCTAAACATGTAATATCATCTATTAATCTGTATGGGAAAAACA 1093
DB 73293 ATTGTTCTTACGCTATGCTAAACATGTAATATCATCTATTAATCTGTATGGGAAAAACA 73352
QY 1094 CGGACCTTAATCTGAAACACCTGGGTAGCGAGACGGGGCGAGAGGGGCTCTGCGAC 1153
DB 73353 CGGACCTTAATCTGAAACACCTGGGTAGCGAGACGGGGCGAGAGGGGCTCTGCGAC 73412
QY 1154 TCAGAGCGAGGCTGAGAGGCGGCGCTCCCTTGCAGAAAGACTGGCATGAGCAGATGGG 1213
DB 73413 TCAGAGCGAGGCTGAGAGGCGGCGCTCCCTTGCAGAAAGACTGGCATGAGCAGATGGG 73472
QY 1214 GACACTCGAGCTGCCCCGCGAGCCTGGGCGGAGCTGCTTAACACTGGGCGCCAGGTGCTG 1273
DB 73473 GACACTCGAGCTGCCCCGCGAGCCTGGGCGGAGCTGCTTAACACTGGGCGCCAGGTGCTG 73532
QY 1274 CAGAAATTAGACCTCGATTAACGTTAACACCACTTCTCATGCTCTTAATGTGTGAT 1333
DB 73533 CAGAAATTAGACCTCGATTAACGTTAACACCACTTCTCATGCTCTTAATGTGTGAT 73592
QY 1334 CCGGCGGCCGAGGGGCTTGTGAGCAGAGGTGCGCTTCAGAGCAGCTCCAGCGACCTT 1393
DB 73593 CCGGCGGCCGAGGGGCTTGTGAGCAGAGGTGCGCTTCAGAGCAGCTCCAGCGACCTT 73652
QY 1394 AAACCTGACCGGCGGACGTCGCGGCGCCGAGGAGAGAAAGAGGCAACCGGAGCCCTCC 1453
DB 73653 AAACCTGACCGGCGGACGTCGCGGCGCCGAGGAGAGAAAGAGGCAACCGGAGCCCTCC 73712
QY 1454 TCAGGCGACAGCCACCTTCAACCACTTCCGTCAGTCGCGACCACTCCCTTCCGCGCTC 1513
DB 73713 TCAGGCGACAGCCACCTTCAACCACTTCCGTCAGTCGCGACCACTCCCTTCCGCGCTC 73772
QY 1514 CCGAGCGCGGCCAGCTGGGGAGCATGCGCATGTCGCGAGCCGGGTTGCCCGGCCACAG 1573
DB 73773 CCGAGCGCGGCCAGCTGGGGAGCATGCGCATGTCGCGAGCCGGGTTGCCCGGCCACAG 73832
QY 1574 CAGGTAGCTGATGCAACTGTGCGGCCCAACCAACCAATCAAGAGACGTTATTTAGCCG 1633
DB 73833 CAGGTAGCTGATGCAACTGTGCGGCCCAACCAACCAATCAAGAGACGTTATTTAGCCG 73892
QY 1634 CCGAGGTGAACATATGCGCAACGCGGCAACCAATCAGAAAGGCGCTTGTGCCGCGAGCCC 1693
DB 73893 CCGAGGTGAACATATGCGCAACGCGGCAACCAATCAGAAAGGCGCTTGTGCCGCGAGCCC 73952
QY 1694 CCGGCGGCCGAGGGGATGTCGCGATGGGTGATGGGTGATGGGTGATGGGTGATGGGTGAT 1753
DB 73953 CCGGCGGCCGAGGGGATGTCGCGATGGGTGATGGGTGATGGGTGATGGGTGATGGGTGAT 74012
QY 1754 CCATCATCCGAGAGGCGCGCGGTTCCCTTGTTCGCGCGGAGCGCGACGCGACG 1813
DB 74013 CCATCATCCGAGAGGCGCGCGGTTCCCTTGTTCGCGCGGAGCGCGACGCGACG 74072
QY 1814 CCGGCACTCGCTAACCGGCGCGCGGCGCGGCGCGCGCGCTGCGTGGGGCGAGGCG 1873
DB 74073 CCGGCACTCGCTAACCGGCGCGCGGCGCGGCGCGCGCGCTGCGTGGGGCGAGGCG 74132
QY 1874 TGGGAGCGGGTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGC 1933
DB 74133 TGGGAGCGGGTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGC 74192
QY 1934 GCGACTAGAGCGCGGAGCGCGGCGGAGCGGAGCGGAGCTGCGTCCGCGCGGAGCTGA 1993
DB 74193 GCGACTAGAGCGCGGAGCGCGGCGGAGCGGAGCGGAGCTGCGTCCGCGCGGAGCTGA 74252
QY 1994 GAAAGATGCTGTCTCTCTGTCTCCACAGCATCTCCATGCGCTACCTGCTGCCAGCGT 2053

DB 74253 AAAAAATGCTGTCTCTCTGTCTCCACAGCATCTCATGCTGCTGCCAGCGT 74312
QY 2054 CGTCTCTCTGGGAGCGGCGGCCACCTTACGTTGTGCTGGGGGCTCGGGGCTGCTC 2113
DB 74313 CGTCTCTCTGGGAGCGGCGGCCACCTTACGTTGTGCTGGGGGCTCGGGGCTGCTC 74372
QY 2114 CGCTTCCTGGCGCGCGCTTCTTCTTCAAGGCTGAGCGAGCGGCTGACTGCTTACCA 2173
DB 74373 CGCTTCCTGGCGCGCGCTTCTTCTTCAAGGCTGAGCGAGCGGCTGACTGCTTACCA 74432
QY 2174 GAGCATGCTGCTCTTCTTCTTCAAGGCTGAGCGAGCGGCTGACTGCTTACCA 2233
DB 74433 GAGCATGCTGCTCTTCTTCTTCAAGGCTGAGCGAGCGGCTGACTGCTTACCA 74492
QY 2234 CCGGCGCTTGGCGCTTCAACCGGAGCTCCCGGGGCGGAGCTCTCGCTCCCCACAGC 2293
DB 74493 CCGGCGCTTGGCGCTTCAACCGGAGCTCCCGGGGCGGAGCTCTCGCTCCCCACAGC 74552
QY 2294 TGGCGAGGCTCACCGGCGGCGCGGCGGAC 2324
DB 74553 TGGCGAGGCTCACCGGCGGCGGCGGCGGAC 74583

RESULT 7
AC091099
LOCUS
DEFINITION
SEQUENCE, 10 unordered pieces.
AC091099.3 GI:22122960
HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 209718)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-782N1
Unpublished
2 (bases 1 to 209718)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Baetien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Camata,J., Campoliano,A., Chang,J., Choepel,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., DeRellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRoque,K.,
Lamazares,R., Landers,T., Lechoczky,U., Levine,R., Liu,G.,
Maclean,C., MacDonald,P., Margulis,N., Matthews,C., Meneus,L.,
McEwan,P., McKernan,K., McPheters,R., Meltrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Riedack,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Roselli,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S.,
Severy,P., Sougnuez,C., Spencer,C., Stange-Thomas,N.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zaloun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (29-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 209718)
Birren,B., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Camata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamet,A.,

Db 127894 CAAGATTAGACTCCGATTAACGTTAAGACCACTTCTCACTGCTTAATTGTGTGAT 127953
 QY 1334 CCGGCGCGCCAGGGCTTGTAGCAGAGGTGGGTTCCAGGCGAGCTCCAGGACCTT 1393
 Db 127954 CCGGCGCGCCAGGGCTTGTAGCAGAGGTGGGTTCCAGGCGAGCTCCAGGACCTT 128013
 QY 1394 AAACGTGACCGGCGCAGCTCCGGCGCCGAGGAGAGACAGAGAGCAGCCGAGCCTTC 1453
 Db 128014 AAACGTGACCGGCGCAGCTCCGGCGCCGAGGAGAGACAGAGAGCAGCCGAGCCTTC 128073
 QY 1454 TCCGGCGCAGCAGCCAGCTTCACTCCAGTTCCGTCACTCCGACCACTCTCCCTCCGCTC 1513
 Db 128074 TCCGGCGCAGCAGCCAGCTTCACTCCAGTTCCGTCACTCCGACCACTCTCCCTCCGCTC 128133
 QY 1514 CGCAGCGCGCGCCAGCTGGGGAGACATGCGCAGTGGCGGAGCGGGGTTGGCCGCGCAGC 1573
 Db 128134 CGCAGCGCGCGCCAGCTGGGGAGACATGCGCAGTGGCGGAGCGGGGTTGGCCGCGCCAG 128193
 QY 1574 CAGGTAGCTGTATCTGCAACTGTCTGGCCCAACCAATCAATCAAGAGACGTGTTATTGCGG 1633
 Db 128194 CAGGTAGCTGTATCTGCAACTGTCTGGCCCAACCAATCAATCAAGAGACGTGTTATTGCGG 128253
 QY 1634 CCGAGGTGGAATCTATGAGCAGCGGCGGACCAATCAAGAGCGCGTTGTTCCCGCGAGCC 1693
 Db 128254 CCGAGGTGGAATCTATGAGCAGCGGCGGACCAATCAAGAGCGCGTTGTTCCCGCGAGCC 128313
 QY 1694 CCGTGGCGCGGAGGAGGATGTGGCGCATGAGGAGTCAAGGAGTGTGAGCATCCCTGAG 1753
 Db 128314 CCGTGGCGCGGAGGAGGATGTGGCGCATGAGGAGTCAAGGAGTGTGAGCATCCCTGAG 128373
 QY 1754 CCATCGATCCGGAGAGGCGCGGAGTTCCCTTCCCGCGGAGCGCGCAGCAGC 1813
 Db 128374 CCATCGATCCGGAGAGGCGCGGAGTTCCCTTCCCGCGGAGCGCGCAGCAGCAGC 128433
 QY 1814 CCGGCACTGCTCTATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 1873
 Db 128434 CCGGCACTGCTCTATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 128493
 QY 1874 TGGGAGCGGGGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1933
 Db 128494 TGGGAGCGGGGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 128553
 QY 1934 GCGACTGAGCGCGCGGAGCGGCGGAGCGCGGAGCGCGGAGCTCGCTGCGCGAGCTGA 1993
 Db 128554 GCGACTGAGCGCGCGGAGCGGCGGAGCGCGGAGCGCGGAGCTCGCTGCGCGAGCTGA 128613
 QY 1994 GAAAGTGTGCTGTCTCCCTGTGTCTCCACAGCTATCTCACTGCTGCTGCTGCTGCT 2053
 Db 128614 AAAGATGTGCTGTCTCCCTGTGTCTCCACAGCTATCTCACTGCTGCTGCTGCTGCT 128673
 QY 2054 CGTGTCTGTGAGCAGCGGCGCGCACTACGTGTGCGCTGGGGGAGTGTGCGGCTGCTC 2113
 Db 128674 CGTGTCTGTGAGCAGCGGCGCGCACTACGTGTGCGCTGGGGGAGTGTGCGGCTGCTC 128733
 QY 2114 GCGCTTCTGCTCCCGCGCGCTTCTACCAAGCGCTGAGCAGCGGCTGTACTGCGTCTACCA 2173
 Db 128734 GCGCTTCTGCTCCCGCGCGCGCTTCTACCAAGCGCTGAGCAGCGGCTGTACTGCGTCTACCA 128793
 QY 2174 GAGCATGTGCTCTTCTTCTTCTGAGATTAACAACCGGGGTCCAGGTAGAGCGGCTCCCGCT 2233
 Db 128794 GAGCATGTGCTCTTCTTCTTCTGAGATTAACAACCGGGGTCCAGGTAGAGCGGCTCCCGCT 128853
 QY 2234 CCGGGGTCTCGGCGTCAACCGGAGCTCCGGGGGCGGAGCTCTCCGCTCCCGCAGC 2293
 Db 128854 CCGGGGTCTCGGCGTCAACCGGAGCTCCGGGGGCGGAGCTCTCCGCTCCCGCAGC 128913
 QY 2294 TGGCGAGGCTCAACCGGCGCGCGCGCGAGC 2324
 Db 128914 TGGCGAGGCTCAACCGGCGCGCGCGCGAGC 128944

RESULT 8
 AC009632/c

LOCUS AC009632 185834 bp DNA linear HTG 07-JUN-2000
 DEFINITION Homo sapiens chromosome 8 clone RP11-397K22 map 8, *** SEQUENCING
 ACCESSION AC009632
 VERSION AC009632.3 GI:8954116
 KEYWORDS HTG; HTGS; PHASEL.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 185834)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 8, clone RP11-397K22
 Unpublished
 2 (bases 1 to 185834)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baker, J., Baldwin, J., Barna, N., Beckert, R., Benn, J., Brown, A.,
 Castele, A., Cerny, J., Colangelo, M., Collins, S., Collamore, A.,
 Cooke, P., Dearrellano, K., Depayre, B., Devon, K., Dewar, K.,
 Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C.,
 Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
 Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
 Karatas, A., Lehoczy, J., Lieu, C., Locke, K., MacDonald, P.,
 Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
 Meidrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
 Naylor, J., Nijhoff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
 Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
 Teefaye, S., Tornella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
 Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
 Direct Submission
 Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 7, 2000 this sequence version replaced gi:5932606.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L2099
 Center clone name: 397_K_22

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 48 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1361: contig of 1361 bp in length
 * 1362 1461: gap of 100 bp
 * 1462 3202: contig of 1741 bp in length
 * 3203 3302: gap of 100 bp
 * 3303 4619: contig of 1317 bp in length
 * 4620 4719: gap of 100 bp
 * 4720 5821: contig of 1102 bp in length
 * 5822 5921: gap of 100 bp
 * 5921 6419: contig of 498 bp in length
 * 6420 6519: gap of 100 bp
 * 6520 8242: contig of 1723 bp in length
 * 8243 8342: gap of 100 bp
 * 8343 9703: contig of 1361 bp in length
 * 9704 9803: gap of 100 bp
 * 9804 10936: contig of 1133 bp in length
 * 10937 11036: gap of 100 bp
 * 11037 12339: contig of 1303 bp in length
 * 12340 12440 12440: gap of 100 bp
 * 12440 13923: contig of 1484 bp in length

Db 104812 GCTACAAAGATGCTTTTATGATTAACAATGTCAGCGCACTACTTAACTACTACA 104753
Qy 914 GAAAAAAGAGGCTCTTTTAAAAAATACAAACCAAGTCCAAAGACTCTGAGAGA 973
Db 104752 GAAAAAAGAGGCTCTTTTAAAAAATACAAACCAAGTCCAAAGACTCTGAGAGA 104693
Qy 974 AATGAAGCAAGATGATTTCTGAAAGGCTTAATAAGATGAGAAATATCTGTTGGG 1033
Db 104652 AATGAAGCAAGATGATTTCTGAAAGGCTTAATAAGATGAGAAATATCTGTTGGG 104633
Qy 1034 ATTGTTCTTACGCTATGATTAACATGTAATTAATCATCATTACTGTATGAGGAAAAACA 1093
Db 104632 ATTGTTCTTACGCTATGATTAACATGTAATTAATCATCATTACTGTATGAGGAAAAACA 104573
Qy 1094 CGGACCTTAATTTCTGAAACCCCTGTGTGCGAGAACGGGCGAGGGGCTGTCTCGAC 1153
Db 104572 CGGACCTTAATTTCTGAAACCCCTGTGTGCGAGAACGGGCGAGGGGCTGTCTCGAC 104513
Qy 1154 TCAGAGCGGAGCTGAGAGAGGCGGCTCCCTTGCAGAAAGACTGTGAGCAGATGGG 1213
Db 104512 TCAGAGCGGAGCTGAGAGAGGCGGCTCCCTTGCAGAAAGACTGTGAGCAGATGGG 104453
Qy 1214 GACACTCGAGCTGCGCCCGGAGCTGCGGCGAGCTGCTCAACCTGTGGCCGAGTGCCTG 1273
Db 104452 GACACTCGAGCTGCGCCCGGAGCTGCGGCGAGCTGCTCAACCTGTGGCCGAGTGCCTG 104393
Qy 1274 CAAGATTTAGACTCCGATTAAGTTAAACAACCACTTTCACATGCTTAATTTGTGAT 1333
Db 104392 CAAGATTTAGACTCCGATTAAGTTAAACAACCACTTTCACATGCTTAATTTGTGAT 104333
Qy 1334 CCGGCGGCTCAGGGGCTTGTGAGCAGAGTGCCTTCCAGAGCACTTCCAGCGACCTT 1393
Db 104332 CCGGCGGCTCAGGGGCTTGTGAGCAGAGTGCCTTCCAGAGCACTTCCAGCGACCTT 104273
Qy 1394 AAACCTGACCGGCGGAGCTGCGGCGGAGGAGCAAAAGAGCAACCCGACCTTCC 1453
Db 104272 AAACCTGACCGGCGGAGCTGCGGCGGAGGAGCAAAAGAGCAACCCGACCTTCC 104213
Qy 1454 TCCGGGCAAGCACCCTTCAACCAAGTTCAGTGCACACCTCCCTTCCGCGCTC 1513
Db 104212 TCCGGGCAAGCACCCTTCAACCAAGTTCAGTGCACACCTCCCTTCCGCGCTC 104153
Qy 1514 CCGAGCGGCGCAGCTGCGGAGGATGCGAGTGCCTGCGAGCGGCTTCCGCGCAG 1573
Db 104152 CCGAGCGGCGCAGCTGCGGAGGATGCGAGTGCCTGCGAGCGGCTTCCGCGCAG 104093
Qy 1574 CAGGTAGCTGTACTGCACTGTGCGCCAAACCAATCAAGAGACTGTATTGCGG 1633
Db 104092 CAGGTAGCTGTACTGCACTGTGCGCCAAACCAATCAAGAGACTGTATTGCGG 104033
Qy 1634 CCGAGGTGGAACCTATGGAACCGGCGCAACATCAGAAAGCGGTTTCCGCGAGCCC 1693
Db 104032 CCGAGGTGGAACCTATGGAACCGGCGCAACATCAGAAAGCGGTTTCCGCGAGCCC 103973
Qy 1694 CCTGCCCGCAGGGGAGATGTGCGCATGAGGTGATGAGGTGATGCAATCCCTGAG 1753
Db 103972 CCTGCCCGCAGGGGAGATGTGCGCATGAGGTGATGAGGTGATGCAATCCCTGAG 103913
Qy 1754 CCAATGATCCGGAGAGGCGCGGCTTCTTGTGCGCGCGGAGCGCGCAGCGC-AG 1812
Db 103912 CCAATGATCCGGAGAGGCGCGGCTTCTTGTGCGCGCGGAGCGCGCAGCGC-AG 103853
Qy 1813 CCGCGGACTCGCTTACCGCGGCGCGGCGCGGCGCGGCGCATGCGGCTGCGGAGCGGAGG 1872
Db 103852 CCGCGGACTCGCTTACCGCGGCGCGGCGCGGCGCGGCGCATGCGGCTGCGGAGCGGAGG 103793
Qy 1873 CTGGGAGCGGAGTGTGCGGCGCGGCGCGGCGCGGCGCGGCGCTGTGCTGTGCG 1932
Db 103792 CTGGGAGCGGAGTGTGCGGCGCGGCGCGGCGCGGCGCGGCGCTGTGCTGTGCG 103733
Qy 1933 CCGGACTGAGGCGCGGAGGCGGCGGAGCGGAGCGGAGCGGAGCTGCTGCGCGAGCTG 1992
Db 103732 CCGGACTGAGGCGCGGAGGCGGCGGAGCGGAGCGGAGCGGAGCTGCTGCGCGAGCTG 103673

Qy 1993 AGAGATGCTGCTGTCCCTGTGTGTCCACAGTACTCCATGCGGTACTGTGCCACG 2052
Db 103672 AAAAGATGCTGCTGTCCCTGTGTGTGTCCACAGTACTCCATGCGGTACTGTGCCACG 103613
Qy 2053 TGTGTCTCTGGGAGCAGGCGCCCACTTACGTGTGTGTGTGTGTGTGTGTGTGTGT 2112
Db 103612 TGTGTCTCTGGGAGCAGGCGCCCACTTACGTGTGTGTGTGTGTGTGTGTGTGTGT 103553
Qy 2113 CCGGCTTCTGCGCGCGCGGCTTCAACAGGCGGTGAGCAGCGGCTGACTGCTTAC 2172
Db 103552 CCGGCTTCTGCGCGCGCGGCTTCAACAGGCGGTGAGCAGCGGCTTACTGCTTAC 103493
Qy 2173 AGAGCATGCTGCTTCTTCTTCCAGATTAACAACCGGAGTTCAGTGAAGCGCTCCGC 2232
Db 103492 AGAGCATGCTGCTTCTTCTTCCAGATTAACAACCGGAGTTCAGTGAAGCGCTCCGC 103433
Qy 2233 TCCGCGGCTTCTGCGGCTTCCAGGCTTCCGCGGCGCGGAGCTTCTGCTTCCCAAG 2232
Db 103432 TCCGCGGCTTCTGCGGCTTCCAGGCTTCCGCGGCGCGGAGCTTCTGCTTCCCAAG 103373
Qy 2293 CTGGCGAGGCTCACCGGCGCGGCGCGGCGGAGC 2324
Db 103372 CTGGCGAGGCTCACCGGCGCGGCGCGGCGGAGC 103341

RESULT 9

AX141275 1509 bp DNA linear PAT 31-MAY-2001

LOCUS AX141275 Sequence 2 from Patent WO0134782.

DEFINITION AX141275

ACCESSION AX141275

VERSION AX141275.1 GI:14281510

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Leung, D.W.

AUTHORS Human lysophosphatidic acid acyltransferase-5(e)

TITLE Patent: WO 0134782-A 2 17-MAY-2001;

JOURNAL CELL THERAPEUTICS, INC. (US)

FEATURES

source

1..1509

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 18.6%; Score 433; DB 6; Length 1509;

Best Local Similarity 99.8%; Pred. No. 4.3e-235;

Matches 483; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1733 TGGGCTGAGCATCTCCCTGAGCCATCCATCCGAGAGGCGCGGTTCCCTTGTGCG 1792
Db 10 TGGGCTGAGCATCTCCCTGAGCCATCCATCCGAGAGGCGCGGTTCCCTTGTGCG 69
Qy 1793 GCCGAGAGGCGGCAACGAGCCCGCACTGCGCTTACCAGGCGCGGCGCGGCGCGCC 1852
Db 70 GCCGAGAGGCGGCAACGAGCCCGCACTGCGCTTACCAGGCGCGGCGCGGCGCGCC 129
Qy 1853 CATGCGGCTGAGGCGGAGGCTGAGGCGGAGGCTGAGGCGGCGGCGCGGCGCGGCGG 1912
Db 130 CATGCGGCTGAGGCGGAGGCTGAGGCGGAGGCTGAGGCGGCGGCGGCGGCGGCGG 189
Qy 1913 GTGATTTGCGCGCTCTCTGCGCGCACTAGAGCCCGGAGAGCGGCGGAGAGCGGAGCGG 1972
Db 190 GTGATTTGCGCGCTCTCTGCGCGCACTAGAGCCCGGAGAGCGGCGGAGAGCGGAGCGG 249
Qy 1973 AGCTGCTGCGCGCGAGCTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2032
Db 250 AGCTGCTGCGCGCGAGCTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309

Qy	2033	GGCGTACCTGCTGGCCACGACGTCGAGCTCTTG3GGAACGGCGGCCACCTACGATGTTGGCCTG	2092
Db	310	GGCGTACCTGCTGGCCACGACGTCGAGCTCTTG3GGAACGGCGGCCACCTACGATGTTGGCCTG	369
Qy	2093	GGGGGCTGGCGGCTGCTCTCCGCTTCTCTGCGCCGCGCTTCTTACCAAGCGCTGAGCA	2152
Db	370	GGGGGCTGGCGGCTGCTCTCCGCTTCTCTGCGCCGCGCTTCTTACCAAGCGCTGAGCA	429
Qy	2153	CCGCGTACCTGCTGCTTACCGAGCATGTCGCTCTTCTTTCGAAATTACACCGGGGT	2212
Db	430	CCGCGTACCTGCTGCTTACCGAGCATGTCGCTCTTCTTTCGAAATTACACCGGGGT	489
Qy	2213	CCAG	2216
Db	490	CCAG	493
RESULT 10	AF375789		
LOCUS	AF375789	1509 bp	mRNA linear PRI 21-MAY-2001
DEFINITION	Homo sapiens lysophosphatidic acid acyltransferase-epsilon		
ACCESSION	AF375789		
VERSION	AF375789.1	GI:14161584	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Leung, D.W.		
TITLE	Cloning and expression of LPAT-epsilon		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1509)		
AUTHORS	Leung, D.W.		
JOURNAL	Direct Submission		
FEATURES	Submitted (02-MAY-2001) Mol. Biol., Cell Therapeutics, Inc., 201 Elliott Ave., W., Suite 400, Seattle, WA 98119, USA		
source	Location/Qualifiers		
	1..1509		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="8"		
	/map="gpter-p23.3"		
	1..1509		
	/gene="LPAT-e"		
	275..1369		
	/gene="LPAT-e"		
	/EC_number="2.3.1.51"		
	/note="1-acylglycerol-3-phosphate O-acyltransferase"		
	/codon_start=1		
	/product="lysophosphatidic acid acyltransferase-epsilon"		
	/protein_id="AAK54809.1"		
	/db_xref="GI:14161585"		
	/translation="MELSLVHTYTSMRYLPSVVLGTAPTYVLAMGVRLLSALFPPR RFYALDRLRCVQOSWVLPFFENYTGVOILLYGLPKNKENIILYLANHSTVDIMIVAF DILALROMALHRYVYLKEGLKMLPLVGCYFAOAGGIYKRSAKENEMENKLOSYYVA DAGPIVLTVEPEGTRVNPBOTKYLASQAFAARGIALVHVLTPRIKATHVAEDCMK KNYDALIYDVTVYVEGDDGCGRRBSPTMTFLCKECKIHIDRIDKQVPEQER MRRVHERFEETDKMLLEFYSPDERKRPFGKSVNSKSIKTLPSMLTISGLTLAG MLMDAGRKLVNMTWITGYTLGLMTWIKK"		
ORIGIN			
Query Match	18.6%	Score 433;	DB 9; Length 1509;
Best Local Similarity	99.8%;	Pred. No. 4.3e-235;	
Matches 483; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1733	TTGGGGTGTGACATCCCTGAGCCATCGATCCGAGAGGCGCGGCTTCCCTTGCTTTGCC	1792
Db	10	TTGGGGTGTGACATCCCTGAGCCATCGATCCGAGAGGCGCGGCTTCCCTTGCTTTGCC	69
Qy	1793	GCCGGAGCGCGCAGCAGCCCCGCACTCGCTACCCGCGCCCGGCGCGCGCGCGGCC	1852

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db	70	GCCTGGAGCGGCGCCACGACGACCCCGACTGCTTACCCGCGCCCGCGGCGCGGCGC	1853	CATCGCGGCTGGGAGCGGAGCGCTGGAGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCG	1912		
Db	130	CATCGCGGCTGGGAGCGGAGCGCTGGAGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCG	130	CATCGCGGCTGGGAGCGGAGCGCTGGAGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCG	189		
Qy	1913	GTGATTGGGCGCGCTGCTGGGCGCGGAGCTGAGGAGCGCGGAGCGCGGAGCGCGGAGCGG	1913	GTGATTGGGCGCGCTGCTGGGCGCGGAGCTGAGGAGCGCGGAGCGCGGAGCGCGGAGCGG	1972		
Qy	1973	AGCTCGGTGCGCGCGGAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2032	AGCTCGGTGCGCGCGGAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2032		
Db	250	AGCTCGGTGCGCGCGGAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	309	AGCTCGGTGCGCGCGGAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	309		
Qy	2033	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2092	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2092		
Db	310	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	369	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	369		
Qy	2093	GGGGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2152	GGGGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2152		
Db	370	GGGGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	429	GGGGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	429		
Qy	2153	CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2212	CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2212		
Db	430	CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	489	CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	489		
Qy	2213	CCAG 2216					
Db	490	CCAG 493					
RESULT 11	BC023550	1503 bp mRNA linear PRI 06-OCT-2003					
LOCUS	BC023550	Homo sapiens acid acyltransferase-epsilon, complete cds.					
ACCESSION	BC023550	NC_19506 IMAGE:412973), complete cds.					
VERSION	BC023550.2	GI:33879276					
KEYWORDS	MGC.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	1	(bases 1 to 1503)					
AUTHORS	Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D., Altschuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marisina, K., Farmer, A.A., Rubin, G.M., Hong, J., Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheer, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loughellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Weller, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.						
TITLE	human and mouse cDNA sequences						
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16899-16903 (2002)					
MEDLINE	22388257						
PUBMED	12477932						
REFERENCE	2	(bases 1 to 1503)					
AUTHORS	Straussberg, R.						
TITLE	Direct Submission						
JOURNAL	Submitted (05-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer						

REMARK
COMMENT

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:23272119.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Hansen, N.L., Grante, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRM Plate: 27 Row: m Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 8922941.
Location/Qualifiers

FEATURES
source

1..1503
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:19506 IMAGE:4129973"
/issue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC 19"
/lab_host="DH10B-R."
/note="Vector: pOTB7"
1..1503
/gene="LPAAT-e"
/note="synonym: FLJ11210"
/db_xref="LocusID:55326"
257..1351
/codon_start=1
/product="LPAAT-e protein"
/protein_id="AAH23550.1"
/db_xref="GI:23272120"
/db_xref="LocusID:55326"
/translation="MLSLVHTYSNRYLPSVVLIGTAPTYVLAMGWRLLSAFLPA
RFOALDDRLVCVOSMVLFFPENYTGVOILLYGDLPKNENIYLANHOSIVDMIVA
DILAIRONLGHRYVYLKEGIKMLPLVGYCPAHOHGIYKRSKFNKEMKRLQSYV
DACTPMYLVTFPEGRYRNPEQTYVLSASQAPAAORGLAVLKVLTPRIKATHVADCM
KNVLDAIVDTVVYEGKDDGQRBSPTMTPEFLCKECPKIHIDRIDKDVPEDEH
MRMLHREFEIKDKMLIEFYESPDPERRRRPFGKSVNSKLSIKTLPSMLILSGLTAG
MLMTDGRKLYVNTWYIGTLGCLMTVITKA"
515..856
/note="PLAC: Region: Phosphate acyltransferases"
/db_xref="CD:smart00563"

ORIGIN

Query Match 18.2%; Score 424; DB 9; Length 1503;
Best Local Similarity 99.8%; Pred. No. 6e-230;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 1 AGCATCCCTAGCATGATCCGGAGAGGCGCGGGTTCCTTGTGCGCGGAGC 1801
1742 AGCATCCCTAGCATGATCCGGAGAGGCGCGGGTTCCTTGTGCGCGGAGC 1801
1802 GGGGAG 1861
DB 61 GGGGAG 120
QY 1862 GGGGAG 1921

DB 121 GGGGAG 180
QY 1922 CGCTGCTGAGCGCGAGTGAAGCGCGGAGCGGAGAGAGAGAGAGAGAGAGAG 1981
DB 181 CGCTGCTGAGCGCGAGTGAAGCGCGGAGCGGAGAGAGAGAGAGAGAGAGAG 240
QY 1982 CGGCGAG 2041
DB 241 CGGCGAG 300
QY 2042 GCTGCCAGAGCGTCTGCTCTGCTGCGGACGCGCCACCTAGTGTGCTGCGGAGCTG 2101
DB 301 GCTGCCAGAGCGTCTGCTCTGCTGCGGACGCGCCACCTAGTGTGCTGCGGAGCTG 360
QY 2102 GCGGCTGCTCTGCGCTCTGCTGCGGAGCGCGGAGCGCGGAGCGCGGAGCTG 2161
DB 361 GCGGCTGCTCTGCGCTCTGCTGCGGAGCGCGGAGCGCGGAGCGCGGAGCTG 420
QY 2162 CTGCGTCTACAG 2216
DB 421 CTGCGTCTACAG 475

RESULT 12
AX268078 3725 bp DNA linear PAT 26-OCT-2001
LOCUS AX268078
DEFINITION Sequence 6 from Patent WO0173051.
ACCESSION AX268078
VERSION AX268078.1 GI:16516593
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1
Meyers, R.A., Rudolph-Owen, L., Macbeth, K.J., Tsai, F.Y. and
Williamson, M.
46743 and 27417, novel human acyltransferase family members and
uses therefor
Patent: NO 0173051-A 6 04-OCT-2001;
JOURNAL
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers

FEATURES
source

1..3725
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
306..1400
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10243.1"
/db_xref="GI:16516594"
/db_xref="REMBL:CAD10243"
/translation="MLSLVHTYSNRYLPSVVLIGTAPTYVLAMGWRLLSAFLPA
RFOALDDRLVCVOSMVLFFPENYTGVOILLYGDLPKNENIYLANHOSIVDMIVA
DILAIRONLGHRYVYLKEGIKMLPLVGYCPAHOHGIYKRSKFNKEMKRLQSYV
DACTPMYLVTFPEGRYRNPEQTYVLSASQAPAAORGLAVLKVLTPRIKATHVADCM
KNVLDAIVDTVVYEGKDDGQRBSPTMTPEFLCKECPKIHIDRIDKDVPEDEH
MRMLHREFEIKDKMLIEFYESPDPERRRRPFGKSVNSKLSIKTLPSMLILSGLTAG
MLMTDGRKLYVNTWYIGTLGCLMTVITKA"

ORIGIN

Query Match 12.1%; Score 281; DB 6; Length 3725;
Best Local Similarity 99.4%; Pred. No. 3e-148;
Matches 501; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
DB 1714 TGGCGATGGGTAGAGGTGATGGGGTGTGAGCATCCCTTAGCATCCGGAGAGGCG 1773
DB 21 TGGCGATGGGTAGAGGTGATGGGGTGTGAGCATCCCTTAGCATCCGGAGAGGCG 80
QY 1774 GGGGTTCCCTTGTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1833
DB 81 GGGGTTCCCTTGTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 140

QY	1834	CCGCGGCGCGCGCGCGCCCAATGCGCGCTGCGGAGCGAGGCTGAGAGCGGGTGGC -GGGCG	1892
Db	141	CCCGGGCGGCGCGCGGCCCATGCGGCGCTGCGGAGCGAGGCTGAGAGCGGGTGGCTGGGCG	200
QY	1893	CGGCGGCGCGGCGCGCGCGGTGAATTGGCGGCTGTGGCCGCACTGAGGCGCGGAGG	1952
Db	201	CGAGAGGCCCGGCGCCCGGCGGGTGAATTGGCGGCTGTGGCCGCACTGAGGCCGCGGAGG	260
QY	1953	CGGCGGCGGAGCGCAAGCGGAGCTGCGCTGCGCGCGAGCTGAGAGATGCTGTCTCTG	2012
Db	261	CGGCGGCGGAGCGCAAGCGGAGCTGCGCTGCGCGCGAGCTGAGAGATGCTGTCTCTG	320
QY	2013	GTGCTCCACAGATTACTCCATGCGGCACTGCGTGGCCAGCGTGGTCTCTGCGGACCGGCG	2072
Db	321	GTGCTCCACAGATTACTCCATGCGGCACTGCGTGGCCAGCGTGGTCTCTGCGGACCGGCG	380
QY	2073	CCCACTTAAGTGTGGCGCTGGGAGGATGTGAGCGGCTGTCTCGGCTTCTCTGCGCGCG	2132
Db	381	CCCACTTAAGTGTGGCGCTGGGAGGATGTGAGCGGCTGTCTCGGCTTCTCTGCGCGCG	440
QY	2133	TTCTACCAAGCGCTGAGCAGACCGGCTGTACTGCGCTACACAGAGATGTGCTCTTCTTC	2192
Db	441	TTCTACCAAGCGCTGAGCAGACCGGCTGTACTGCGCTACACAGAGATGTGCTCTTCTTC	500
QY	2193	TTGAGGAATTAACCGGGGTCCAG	2216
Db	501	TTGAGGAATTAACCGGGGTCCAG	524

RESULT 13				
BD196515				
LOCUS	BD196515	4558 bp	DNA	linear
DEFINITION	Prostateal cancer gene.			
ACCESSION	BD196515			
VERSION	BD196515.1	GI:33006285		
KEYWORDS	JP 2002516657-A/104.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Euharyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 4558)			
TITLE	Cohen,D., Blumenfeld,M., Chumakov,I. and Bougueleret,L.			
JOURNAL	Prostatic cancer gene			
COMMENT	Parent: JP 2002516657-A 104 11-JUN-2002;			
	GENSET			
	OS Homo sapiens (human)			
	PN JP 2002516657-A/104			
	PD 11-JUN-2002			
	PF 22-DEC-1998 JP 2000525562			
	PR 22-DEC-1997 US 08/966306 09-SEP-1998 US 60/099668 PI			
	DANIEL COHEN, MARIA BLUMENFELD, ILYA CHUMAKOV, LIDIE BOUGUELERET			
	PC C12N15/09, C12N15/09, A01K67/027, C07K14/47, C07K16/18, C12N1/15, PC			
	C12N1/19,			
	PC C12N1/21, C12N5/10, C12N5/10, C12P21/08, C12Q1/68, G01N33/50 PC			
	PC C12N15/00, C12N5/00,			
	CC C12N5/00, C12N15/00			
	CC ATG			
	CC TGA			
	CC AATATA			
	PH Key	Location/Qualifiers		
	FT misc feature	31..33		
	FT misc feature	235..237		
	FT polyA_signal	4531..4536.		
FEATURES	Location/Qualifiers			
Source	1..4558			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	9.3%	Score 217;	DB 6;	Length 4558;
Best Local Similarity	100.0%;	Pred. No. 1,1e-111;		

	Matches	211,	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	2001	CTGCTGTCCCTGGTGTGCTCCACAGATTCATGAGCGTCACTGCTGCTGCCACAGCTGTGTCTC	2050							
Db	1	CTGTGTGTCCTTGGTGTCTCCACAGATTCATGAGCGTCACTGCTGCTGCCACAGCTGTGTCTC	60							
QY	2061	CTGGGACACGGCGCCACCTACGTTGTGGCTGGGGGGGTCTGTGCGGCTCTCCGCGCTTC	2120							
Db	61	CTGGGACACGGCGCCACCTACGTTGTGGCTGGGGGGGTCTGTGCGGCTCTCTCGGCTTC	120							
QY	2121	CTGCGCGCGCGCTTTCACCAAGCGGTGGACGACCGGCTGTACTGTGCTTACAGAGCATG	2180							
Db	121	CTGCGCGCGCGCTTTCACCAAGCGGTGGACGACCGGCTGTACTGTGCTTACAGAGCATG	180							
QY	2181	GTGCTCTTCTTCTTCGAGATTACACCGGGGTCAGG	2217							
Db	181	GTGCTCTTCTTCTTCGAGATTACACCGGGGTCAGG	217							

LOCUS	BD196497	216 bp	DNA	linear	PAT 17-JUN-2003
DEFINITION	Prostatic cancer gene.				
ACCESSION	BD196497				
VERSION	BD196497.1	GI:33006267			
KEYWORDS	JP 2002516657-A/86.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 216)				
TITLE	Cohen,D., Blumenfeld,M., Chumakov,I. and Bougueleret,L.				
JOURNAL	Prostatic cancer gene				
COMMENT	Patent: JP 2002516657-A 86 11-JUN-2002;				
	GENSER				
	OS Homo sapiens (human)				
	PN JP 2002516657-A/86				
	PD 11-JUN-2002				
	PF 22-DEC-1998 JP 2000525562				
	PIR 22-DEC-1997 US 08/996356 09-SEP-1998 US 60/099658 PI				
	DANIEL COHEN,MARTA BLUMENFELD,IYA CHUMAKOV,IYDIE BOUGUELERET PC				
	C12N5/09,C12N5/09,A01K67/027,C07K14/47,C07K16/18,C12N1/15, PC				
	C12N1/19				
	PC C12N1/21,C12N5/10,C12N5/10,C12P21/08,C12O1/68,G01N3/50 PC				
	,C12N5/00,C12N5/00,				
	PC C12N5/00,C12N5/00				
	CC Prostatic cancer gene				
	PH Key	location/Qualifiers			
	FT source	1..216			
		location/Qualifiers			
		FT			
FEATURES					
source	location/Qualifiers				
	1..216				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	9.3%; Score 216; DB 6; Length 216;				
Best Local Similarity	100.0%; Pred. No. 3e-111;				
Matches 216; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	2001 CTGCTGTCCCTGTGTGTCACACACGTA				
Db	1 CTGTGTCCCTGTGTGTCACACACGTA				
QY	2061 CTGGGACACGGGCGCCACCTACAGT				
Db	61 CTGGGACACGGGCGCCACCTACAGT				
QY	2121 CTGGCGGCGCGCTTCTACCAAGGCG				
Db	121 CTGGCGGCGCGCTTCTACCAAGGCG				

QY 2181 GNGCTCTTCTCTGAGATTACACGGGGTCAG 2216
|||||
Db 181 GTGCTCTTCTTCGAGATTACACGGGGTCAG 216

RESULT 15

BD196514 4682 bp DNA linear PAT 17-JUL-2003
DEFINITION BD196514
ACCESSION BD196514.1 GI:33006284
VERSION JP 2002516657-A/103.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4682)
AUTHORS Cohen, D., Blumenfeld, M., Chumakov, I. and Bougueleret, L.
TITLE Prostatic cancer gene
JOURNAL Patent: JP 2002516657-A 103 11-JUN-2002;
GENSET

COMMENT OS Homo sapiens (human)
PN JP 2002516657-A/103
PD 11-JUN-2002
PE 22-DEC-1998 JP 200525562
PR 22-DEC-1997 US 08/996306 09-SRP-1998 US 60/099658 PI
DANIEL COHEN, MARTA BLUMENFELD, ILVA CHUMAKOV, LYDIE BOUGUELERET
C12N15/09, C12N15/09, A01K67/027, C07K14/47, C07K16/18, C12N1/15, PC
C12N1/19,
PC C12N1/21, C12N5/10, C12N5/10, C12P21/08, C12Q1/68, G01N33/50 PC
, C12N15/00, C12N5/00,
PC C12N5/00, C12N15/00
CC ATG
CC TGA
CC AATAAA
FH Key location/Qualifiers
FT misc feature 31..33
FT misc feature 301..303
FT polyA_signal 4655..4660.
location/Qualifiers
1..4682
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
ORIGIN

Query Match 9.3%; Score 216; DB 6; Length 4682;
Best Local Similarity 100.0%; Pred. No. 4e-111;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2001 CTGCTGTCCCTGGTGTCCACAGTACTCCATGCGCTACCTGTGCGCAGCGTGTGCTC 2060
|||||
Db 1 CTGCTGTCCCTGGTGTCCACAGTACTCCATGCGCTACCTGTGCGCAGCGTGTGCTC 60
QY 2061 CTGCGGACCGCGCCACCTACGTGTGCGCTGCGGGGCTGCGGCTGCTCTCCGCTTC 2120
|||||
Db 61 CTGCGGACCGCGCCACCTACGTGTGCGCTGCGGGGCTGCGGCTGCTCTCCGCTTC 120
QY 2121 CTGCGCGCGCGCTTCTACCAAGCGCTGAGAGACCGGCTGTACTGCGTCTTACCAAGACATG 2180
|||||
Db 121 CTGCGCGCGCGCTTCTACCAAGCGCTGAGAGACCGGCTGTACTGCGTCTTACCAAGACATG 180
QY 2181 GTGCTCTTCTTCTGAGATTACACCGGGGTCAG 2216
|||||
Db 181 GTGCTCTTCTTCTGAGATTACACCGGGGTCAG 216

Search completed: May 26, 2004, 10:29:32
Job time : 8839.79 secs

Paraday Avenue Genoscope sequence ID : CS0DC010CE04QP1.

FEATURES	Location/Qualifiers
source	1. .1201

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC01.0VY107"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone="1st-Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="rib strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	19.6%	Score 455;	DB 13;	Length 1201;
Best Local Similarity	99.8%;	Pred. No. 2.6e-207;		
Matches 505;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1711	ATGTGGCGATGGGTGGAGGGTCATGGGGGTGTAGCATCCCTAGGCATCGATCCCGGAGGG	1770
Db	86	ATGTGGCGATGGGTGGAGGGTCATGGGGGTGTAGCATCCCTAGGCATCGATCCCGGAGGG	145
QY	1771	CCGCGGGTTCCTTGTCTTTGTCGCGCCGCGAGCGCGACGACGAGCCCGACCTGCTTACC	1830
Db	146	CCGCGGGTTCCTTGTCTTTGTCGCGCGAGCGCGACGACGAGCCCGACCTGCTTACC	205
QY	1831	GGCCCGGGGCGGGCGCGCGCCCAATGACGGCTGGGGGCGAGAGCTGGAGCGCGGTGGCGGG	1890
Db	206	GGCCCGGGGCGGGCGCGCGCCCAATGACGGCTGGGGGCGAGAGCTGGAGCGCGGTGGCGGG	265
QY	1891	CGCGGCGGGCCCGGGGCCGGGCGGTGATTTGCGCGCTGTGGCCCGGACTGAGGCCCGGGA	1950
Db	266	CGCGGCGGGCCCGGGGCCGGGCGGTGATTTGCGCGCTGTGGCCCGGACTGAGGGCCCGGGA	325
QY	1951	GGCGGGGCGGGGAGCGGACGGAGCTGCTGCGCGCGAGCTGAGAGAAATGTGCTGTCC	2010
Db	326	GGCGGGGCGGGGAGCGGACGGAGCTGCTGCGCGCGAGCTGAGAGAAATGTGCTGTCC	385
QY	2011	TGTTGCTGCACACGTACTCTCATGCGGTACCTGTCCCGACGTCGTGCTCTTGGGACCG	2070
Db	386	TGTTGCTGCACACGTACTCTCATGCGGTACCTGTCCCGACGTCGTGCTCTTGGGACCG	445
QY	2071	CGCCACCTACGTGTGTGGCTGTGGGGGGGTCTGAGCGCGCTCTCCGACCTTCTGTGCCGCGC	2130
Db	446	CGCCACCTACGTGTGTGGCTGTGGGGGGGTCTGAGCGCGCTCTCTCCGACCTTCTGTGCCGCGC	505
QY	2131	GCTTCTACCAAGCGCTGGACGACCGGCTGTACTGTGCTTACCAAGACATGTGTCTTCT	2190
Db	506	GCTTCTACCAAGCGCTGGACGACCGGCTGTACTGTGCTTACCAAGACATGTGTCTTCT	565
QY	2191	TCTTGAGAAATTACACGGGGGTCCAG	2216
Db	566	TCTTGAGAAATTACACGGGGGTCCAG	591

```

FEATURES
source
    Email: c9abbs-remail.nih.gov
    Tissue Procurement: ATCC
    CDNA Library Preparation: Ling Hong/Rubin Laboratory
    CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov
    Plate: L16CM1774 row: f column: 05
    High quality sequence stop: 753.
    Location/Qualifiers
        1..822
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4892932"
        /tissue_type="neuroblastoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH MGC 19"
        /note="Organ: brain; Vector: pOT81; Site 1: XhoI; Site 2:
        EcoRI; cDNA made by oligo-dT priming. Directionally
        cloned into EcoRI/XhoI sites using the following 5'
        adaptor: GGCACGAG(G). Library constructed by Ling Hong
        in the laboratory of Gerald M. Rubin (University of
        California, Berkeley) using ZAP-cDNA synthesis kit
        (Stratagene) and Superscript II RT (Life Technologies).
        Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match	18.0%;	Score 418;	DB 12;	length 822;
Best Local Similarity	99.8%;	Pred. No. 1.7e-189;		
Matches 468;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

QY	1748	CTCTAAGCATGATCTCGAGAGAGGCGGGAGGTTCCCTTGCTTTGCGCGCGAGGCGGCA	1807
Db	7	CTTAAAGCAATGATCTCGAGAGAGCGCGAGTTCCCTTGCTTTGCGCGCGAGGCGGCA	66
QY	1808	CGACGCCCGCACTTCGCTTACCCCGAGCCCGAGCGAGCGCGGCAATGCGGCTGGAGGC	1867
Db	67	CGAGACCCCGCACTTCGCTTACCCCGAGCCCGAGCGAGCGCGGCAATGCGGCTGGAGGC	126
QY	1868	GGAGGCTGGAGCGGGGTGGCGGGGCGGCGGCGCGGCGCGGGCGGGGATTTGGGCGGCTG	1927
Db	127	GGAGGCTGGAGCGGGGTGGCGGGGCGGCGGCGCGGCGCGGGCGGGGATTTGGGCGGCTG	186
QY	1928	CTGGCCCGGACTGAGAGCCCGGAGGCGGAGCGAGGAGCGACAGCGAGCTCGCTCCGCGG	1987
Db	187	CTGGCCCGGACTGAGAGCCCGGAGGCGGGGAGGAGGAGGAGCGAGCGGAGCTCGCTCCGCGG	246
QY	1988	AGCTGAGAGATGCTGCTGTCTCTCTGTTGCTTCAACGTACTTCATTCGGCTTACTGCTGGC	2047
Db	247	AGCTGAGAGATGCTGCTGTCTCTCTGTTGCTTCAACAGTAATTCATGGGCTTACTGCTGGC	306
QY	2048	CAGGTCGTGTCTCTTGGGACAGGAGCCACCTACGTGTGGCTCGGGGAGGCTGAGCGAGCT	2107
Db	307	CAGGTCGTGTCTCTTGGGACAGGAGCCACCTACGTGTGGCTCTGGGGAGGCTTGGCGGCT	366
QY	2108	GCTCTCGGCTTCTGTCGCGCGGCGGCTTTCACAAAGCGTGGAGCAACCGGCTGTACTGCGT	2167
Db	367	GCTCTCGGCTTCTGTCGCGCGGCGGCTTTCACAAAGCGTGGAGCAACCGGCTGTACTGCGT	426
QY	2168	CTACCAAGCATGTGTGCTCTTCTTCTTGANAATTACACGGGGTTCAG	2216
Db	427	CTACCAAGCATGTGTGCTCTTCTTCTTCTTGANAATTACACGGGGTTCAG	475

RESULT 2				
LOCUS	BI197695	822 bp	mRNA	linear
DEFINITION	602757619p1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:4892932 5', _			
ACCESSION	BI197695			
VERSION	BI197695.1			
KEYWORDS	EST, GI:14652715			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Eumaliota; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 822)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1995)			
COMMENT	Contact: Robert Strausberg, Ph.D.			

RESULT 3
AL560635
LOCUS
DEFINITION
Accession
Version
Keywords
1005 bp mRNA linear EST 31-MAY-2001
AL560635 Homo sapiens B cells (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL003YD01 5'-PRIME, mRNA sequence.
AL560635
AL560635.2 GI:31284765
EST.

SOURCE	ORGANISM
Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1005)	Li, W.B., Gruber, C., Jessee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001)
JOURNAL COMMENT	On Feb 15, 2001 this sequence version replaced gi:112907288. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr, Web : ww.genoscope.cns.fr Library was constructed by life technologists, a division of Invitrogen. This sequence belongs to sequence cluster 5026.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSDDL003CB01QPLcluster=5026.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen.Faraday Avenue Genoscope sequence ID : CSDDL003CB01QPL.
FEATURES	Location/Qualifiers
Source	1..1005 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSDDL003YD01" /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED" /cell_line="RAMOS CELL LINE". /clone_id="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
ORIGIN	/note="1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
. Query Match	15.9%; Score 369; DB 9; Length 1005;
Best Local Similarity	99.4%; Pred. No. 6,4e-166;
Matches 519; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Oy	1803 GCGCAGCAGCCCGCACTGCGTACCACGGGCCGGCGCGCGGCATGCGCTG 1862
Db	157 GCGCAGCAGCCCGCACTGCGTACCACGGGCCGGCGCGCGGCATGCGCTG 216
Oy	1863 GGGCGGAGGCTGGAGAGGGGTGGGGCGCGCGCGGCCCGCGGCGTGATTGGCC 1922
Db	217 GGGCGGAGGCTGGAGAGGGGTGGGGCGCGCGCGGCCCGCGGCGTGATTGGCC 276
Oy	1923 GCTGCTGGCGCGCATGAGGCCCGGAGAGCGGGCGGGGAGCGCAGCGGACTGCTGC 1982
Db	277 GCTGCTGGCGCGCATGAGGCCCGGAGAGCGGGCGGGGAGCGCAGCGGACTGCTGC 336
Oy	1983 CGCGAGCTGAAGAATCTGCTGTCCCTGTGCTCTACAAGTAATCCATGCGCTA 2042
Db	337 CGCGAGCTGAAGAATCTGCTGTCCCTGTGCTCTACAAGTAATCCATGCGCTA 396
Oy	2043 CTGGCACAAGTCGTGCTCTGCGGACCGGCGCACATAAGTTGGTGGCGGGGGCTCG 2102
Db	397 CTGGCACAAGTCGTGCTCTGCGGACCGGCGCACATAAGTTGGTGGCGGGGGCTCG 456
Oy	2103 CGGCTGCTCTCGGCTTCTGCGCGCGCGCTTCTACCAAGGCTGGAAGCAGCGCTGAC 2162
Db	457 CGGCTGCTCTCGGCTTCTGCGCGCGCGCTTCTACCAAGGCTGGAAGCAGCGCTGAC 516
Oy	2163 TGGCTTACCAAGATGCTCTTTCTTTCAGAAATTACACCGGGTCAAGTGAAC 2222
Db	517 TGGCTTACCAAGATGCTCTTTCTTTCAGAAATTACACCGGGTCAAGTGAAC 576
Oy	2223 CGCGTCCGCGTCCGGGATCTGGAGCTCCACCGGACTCCGGGAGCGGAGACTTCGCG 2282
Db	577 CGCGTCCGCGTCCGGGATCTGGAGCTCCACCGGACTCCGGGAGCGGAGACTTCGCG 636

Oy		2283	TCCCCCAGACTGGCGAGGCTGCACCCGCCGCGGCCGCGAAC	2324
Dd		637	TCCCCCACTGTGCGAGGTCACCCGCCGCCGCCGCGAAC	678
RESULT 4	Bt913991	787 bp	mRNA	linear EST 16-OCT-2001
LOCUS	Bt913991			
DEFINITION				
	mRNA sequence.			
ACCESSION	Bt913991			
VERSION	Bt913991.1	GI:16178288		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 787) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)			
AUTHORS	Contact: Robert Strausberg, Ph.D. Email: cgapds-room@nih.nh.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc.			
TITLE	Genomic Arrayed by: The I.M.A.G.E. Consortium (LMNL)			
JOURNAL	DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNl at: http://image.llnl.gov Plate: LRAM11617 row: 1 column: 18			
COMMENT	High quality sequence stop: 700. Location/Qualifiers . . . 787			

Query Match	14.4%	Score 335	DB 12	Length 787
Best Local Similarity	100.0%	Prod. No. 1.5e-149		
Matches 335	Conservative 0	Mismatches 0	Indels 0	Gaps 0
1990	CTGAGAGAGTCTGCTGTGCTCTGTGAGTCTCAACAGTACTTCATGGCTACCTGTGACCA	2045		
5	CTGAGAGAGTCTGCTGTGCTCTGTGAGTCTCAACAGTACTTCATGGCTACCTGTGACCA	64		
2050	GCGTGTGTCTCTTGGGCAAGGCGCCACCTACGTTTGGCTGGGGGGTCTGGCGGGCTGC	2105		
65	GCGTGTGTCTCTTGGGCAAGGCGCCACCTACGTTTGGCTGGGGGGTCTGGCGGGCTGC	124		
2110	TCTCCGCTCTCTGCGCGCGCGCTTCAACAAAGGCTGGACGACCGGCTGACTGCGCT	2165		
125	TCTCCGCTCTCTGCGCGCGCGCTTCAACAAAGGCTGGACGACCGGCTGACTGCGCT	184		
2170	ACCAAGAGATGTCTCTTCTTCTTCAGAAATTACACCGGGGTCAGGTGAGCGGCTCC	2225		
185	ACCAAGAGATGTCTCTTCTTCTTCAGAAATTACACCGGGGTCAGGTGAGCGGCTCC	244		
2230	GCGTCCGGGCTCTGGGCTCAACCGGAGCTCCGGGGGCGGAGCTTCCGCTTCCCA	2285		

```
Db 245 GCGTCCCGGCTCCGCGCCGACCCGAGCTCCGGGGGGGCGGAGACTCTCCGCTCCCA 304
QY 2290 CAGCTGGCGAGGGTCAACCCGCGCGCGCGGAC 2324
Db 305 CAGCTGGCGAGGGTCAACCCGCGCGCGCGGAC 339

RESULT 5
LOCUS B116158 663 bp mRNA linear EST 26-JUN-2001
DEFINITION 602866369F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5015128 5',
mRNA sequence.
ACCESSION B116158
VERSION B116158.1 GI:14567059
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 663)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNCM1823 row: 1 column: 17
High quality sequence stop: 517.
Location/Qualifiers
1..663
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5015128"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```
Query Match 14.2%; Score 330; DB 12; Length 663;
Best Local Similarity 99.6%; Pred. No. 3.8e-147;
Matches 500; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

```
QY 1776 GGGATGGGAGGATCGGAGTGGAGATCCCTGAGCATCGCGGAGGCGCGG 1775
Db 1 GGAATGGGAGGATCGGAGTGGAGATCCCTGAGCATCGCGGAGGCGCGG 60
QY 1776 GGTTCCTTGTCTTTCGCGCGGAGGCGGAGGCGCGGAGGCGCGGCGG 1835
Db 61 GGTTCCTTGTCTTTCGCGCGGAGGCGGAGGCGCGGAGGCGCGGCGG 120
QY 1836 CGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1894
Db 121 CGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 1895 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1954
Db 181 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
```

```
QY 1955 GGGCGGAGCGGAGCGGAGCTCGCTGCGCGCGAGCTGAGAAGATGCTGCTCCGT 2014
Db 241 GGGCGGAGCGGAGCGGAGCTCGCTGCGCGCGAGCTGAGAAGATGCTGCTCCGT 300
QY 2015 GCTCCACAGTACTCATGCGCTACTGCTGCGCGAGCGCTGCTCTCGGCGAGCGGCGC 2074
Db 301 GCTCCACAGTACTCATGCGCTACTGCTGCGCGAGCGCTGCTCTCGGCGAGCGGCGC 360
QY 2075 CACCTACGTTGTGGCTGGGGGGGTGGCGGCTGCTCCGCGCTTCCCGCGCGGCTT 2134
Db 361 CACCTACGTTGTGGCTGGGGGGGTGGCGGCTGCTCCGCGCTTCCCGCGCGGCTT 420
QY 2135 CTACCAAGCGCTGAGCGAGCGGCTGTAAGTGTCTACAGAGCATGAGTCTTCTT 2194
Db 421 CTACCAAGCGCTGAGCGAGCGGCTGTAAGTGTCTACAGAGCATGAGTCTTCTT 480
QY 2195 CGAGATTACACCGGCGGTCAG 2216
Db 481 CGAGATTACACCGGCGGTCAG 502

RESULT 6
LOCUS BG828889 717 bp mRNA linear EST 22-MAY-2001
DEFINITION 602752872F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4905832 5',
mRNA sequence.
ACCESSION BG828889
VERSION BG828889.1 GI:14176476
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 717)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM1807 row: 0 column: 17
High quality sequence stop: 681.
Location/Qualifiers
1..717
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4905832"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

FEATURES

source

ORIGIN

```
Query Match 12.5%; Score 291; DB 12; Length 717;
Best Local Similarity 99.8%; Pred. No. 2.3e-128;
Matches 411; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1747 CCTGAGCATGATCCGAGGAGGCGCGGCTCTTGTTCCTTTCGCGCGGAGCGGCG 1806
```

Db	2	CCCTGAGGCATCGATCCGGAGAGGCCGGGGGTTCCCTTGGCTTTGGCCGCGGAGACGGGCGC	61
QY	1807	ACGACAGCCCCGCACTGGCTTACCCGAGCCCGGAGCGGCGGCGGCCCATATGCGGCTTGGAGG	1866
Db	62	ACGAGAGCCCCGCACTGGCTTACCCGAGCCCGGGGCGGCGGCGGCCCATATGGGCTGGGGG	121
QY	1867	CGAGAGCTGGAGCGGGGTGGCGGGCCGCGGCGGCCCGGAGCCGATGATTGGCCGGCT	1926
Db	122	CGAGAGCTGGAGCGGGGTGGCGGGCCGCGGCGGCCCGGAGCCGATGATTGGCCGGCT	180
QY	1927	GCTGGCGCGCACTAGAGCCCGGGAGGCGGGAGCGGAGGCTGGCTGGCGGC	1986
Db	181	GCTGGCGCGCACTAGAGGCCCGGGAGGCGGGAGGAGGCGGAGCTGGCTGGCGGC	240
QY	1987	GAGCTGAGGAATGCTGCTGTCCTTGCTGCTCCACAGTACTCCATGGGCTACTGCTGC	2046
Db	241	GAGCTGAGGAATGCTGCTGTCCTTGCTGCTCCACAGTACTCCATGGGCTACTGCTGC	300
QY	2047	CCAGGTCGTCCTCTGGGCAAGGCGGCCCATCTACGTGTTGGCCCTGGGGGGGTCGGGCGC	2106
Db	301	CCAGGTCGTCCTCTGGGCAAGGCGGCCCATCTACGTGTTGGCCCTGGGGGGGTCGGGCGC	360
QY	2107	TGCTCTCGAGCTTCTGTCGCGCGCGGCTTCTTACAAAGCGCTGAAGACCGGCT	2158
Db	361	TGCTCTCGAGCTTCTGTCGCGCGCGGCTTCTTACAAAGCGCTGAAGACCGGCT	412

RESULT 7	BU429790	LOCUS	DEFINITION
	BU429790	313 bp	mRNA linear EST 09-SEP-2002
		UI-HF-BNO-aff-f-05-0-UI.r1	NIH MGC_50 Homo sapiens cDNA clone IMAGE:3066752 5', mRNA sequence.

VERSION BU429790.1 GI:22768277
KEYWORDS EST.

SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE	1 (bases 1 to 313)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Eco RI site shown at the beginning of the sequence
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be

Found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html

The following repetitive elements were found in this cDNA sequence:

```
110-233, >GC_rich#Low complexity
Seq primer: M13 Forward.
```

FEATURES	Location/Qualifiers
source	1. .313

```
/organism="Homo sapiens"  
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"  
/clone="IMAGE:3066752"
```

```
/tissue_type="lymph"  
/cell_type="germinal center B cells"
```

```
/cell_line="MGC85"  
/lab_host="DH10B (LTI)"
```

```

/clone_1b="NIH_MGC_50"
/notes="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI,
```

Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by

Louis M. Staudt, Ph.D. Library preparation by Maria de

ORIGIN

Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D.

Query Match	12.5%	Score 290;	DB 13;	Length 313;
Best Local Similarity	100.0%	Pred. No. 7e-128;		
Matches 290;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1687 GGAGCCCCCTGCCCCGGCAGGGGGATGTGGCAGATGGGTGGGGGCATGGGGTGTGAGCCT 1746

Db 7 GGAGCCCCCTGCCCCGGCAGGGGGATGTGGCAGATGGGTGTGAGGCATGGGGTGTGAGCCT 66

QY	1747	CCCTGAGCATCGATCCGGAGGGCCGGGGTTCCTTGCTTGGCCGCCGGAGCCGCC	1808
Db	67	CCCTGAGCATCATCCGGAGGGCCGGGGTTCCTTGCTTGGCCGCCGGAGCCGCC	126

QY	1807	ACGCAGCCCCCGCACTCGCTACCCCGCCCGCGCGCGCGCGCCATCGCGCTGGGCG	1866
Db	127	ACGCAGCCCCCGCACTCGCTACCCCGCCCGCGCGCGCGCGCCATCGCGCTGGGCG	186

Qy	1867	CGAGGCTGGAGCGGGTGGCGGGCGCGCGCGCCCGGCGGTGATTGGCGGCT	1922
Db	187	CGAGGCTGGAGCGGGTGGCGGGCGCGCGCGCCCGGCGGTGATTGGCGGCT	246

Qy	1927	GCTGGCCCGACTGAGGCCCGGAGGCCGGGCGAGCGCGAGCT	1976
Db	247	GCTGGCCCGACTGAGGCCCGGAGGCCGGGCGAGCGCGAGCT	296

RESULT 8
AW969608/c

LOCUS	644 bp	mrna	linear	EST 01-JUN-2000
DEFINITION	EST361685	MAGE	resequeences	
ACCESSION	AW969608			
VERSION	AW969608.1	GI:8159452		
				mrna sequence.

ORGANISM
KEYWORDS
EST.
SOURCE
Homo sapiens
Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

REFERENCE
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 644)
Hedge, P., Qi, R., Abernathy, K., Dhary, S., Gaspard, R., Gay, C.,
Holt, J. P., Saeed, A., Sharrow, V., Lee, N. H., Yeaman, T. J.,
and Holt, J. P.

TITLE Quackenbush, J. Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray

COMMENT
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA

Fax: 301 838 0208
Email: johnq@igr.org
Plate: 267

```
FEATURES location/Qualifiers
source 1..644
/organism="Homo sapiens"
```

```

//db_xref="taxon:9606"
/clone_lib="MAGE_resesources, MAGE"
/note="vector: pBluescriptSKm"

```

Query Match	12.1%	Score 281; DB 10; Length 644;
Best Local Similarity	99.7%	Pred. No. 1.5e-123;

	MATCHES	531	CONSERVATIVE	0	MISMATCHES	17	INDELS	0	GAPS	0
QY		658	AATATAATATAATATAATATAATATAGGTAAAGCATTCCGCAATCTTTTAAAGT							717

718 ACAACTATCTCCTTGAAGGGTTACATATTAACCATTTTACCA CAGCCAAAGGGAGGA 777
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

Db	444	ACAACTATCTCGTTGAAAAGGGTTACATATTTAAACATTTTTCACACAGCAAAAGGGAGGA	385
Qy	778	GAAAGATCCAAAAGCTCTTGATGTGCTTTAACTCAATATAAACAGTTATCCACCTTC	837
Db	384	GAAGAATCCAAAAGTCTCTGATCTGCTTTAACTCAATATAAACAGTTATCCACCTTT	325
Qy	838	GTAGCTTTAGTGAAGGCTACAAAAGTATGCTTTATGATTCACATGTGCACGCAAC	897
Db	324	GTAGCTTTTAGTGAAGGCTACAAAAGTATGCTTTTATGATTCACATGTGCACGCAAC	265
Qy	898	TACTTAAATTACTACAGAAAAAAACGAGGCTCTTATTTAAAAAAAATCAGAAACAATC	957
Db	264	TACTTAAATTACTACAGAAAAAAACGAGGCTCTTATTTAAAAAAAATCAGAAACAATC	205
Qy	958	CAACAGACTCTGAGGAAATGAAACCAAGATGA	989
Db	204	CAACAGACTCTGAGGAAATGAAACCAAGATGA	173

RESULT 9	AL518744	1201 bp	mRNA	linear	EST 12-MAY-2003
LOCUS	AL518744				
DEFINITION	AL518744 Homo sapiens NEUROBLASTOMA Homo sapiens CDNA clone CSODAO11YG14 5-PRIME, mRNA sequence.				
ACCESSION	AL518744				
VERSION	AL518744.2	GI:30537349			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Li W.B., Gruber,C., Jessee,J. and Potayes,D.				
AUTHORS	Full-length cDNA libraries and normalization				
TITLE	Unpublished (2001)				
COMMENT	On Feb 13, 2001 this sequence version replaced gi:12782237.				

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segrete@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5026.f. For
more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DA011BD07QP1&cluster=5026.f>. Contact :
Feng Liang Email : fliang@lifetech.com URL : <http://www.invitrogen.com/> Invitrogen Corporation 1600
Paradise Avenue genoscope sequence ID : CS0DA011BD07QP1.

```

FEATURES
source
    location/Qualifiers
1..1201
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CSODA011Y14"
    /tissue_type="NEUROBLASTOMA"
    /clone_lib="Homo sapiens NEUROBLASTOMA"
    /note="Vector: pcwvSPORT 5: 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pcwvSPORT 6 vector.
Library was not normalized."

```

	Query Match	11.2%	Score 261;	DB 9;	Length 1201;
	Best Local Similarity	99.3%;	Pred. No. 6.1e-114;	Mismatches 411; Conservative	0;
				Indels 3;	Gaps 0;
Oy	1803	GGCAGCGAGGCCGCACTGCCTTCCCGGGCCCCGGGCGGGCGAGCGCGCATGGAGCGTG	1867		
Dd	173	GGCAGCGAGGCCGCACTGCCTTCCCGGGCCCCGGGCGGGCGAGCGCGCATGGAGCGTG	232		
Oy	1863	GGGCGGAGGCTGTGGAGCGGGGTGGCGGGCGCGCGCGCCCCTGGGCGGGGTGAATTGGCC	1922		

Db	233	GGGGCGAAGGTGGAGACGGGTGGCGGGGCGCGCCCGGAGCCCGGGAGCGTATGGAC	292
QY	1923	GCTCTGCTGCTCCGCACTGAGAGCCCGGGAGAGCGGGGAGCCGAGAGCGTACGTG	1982
Db	293	GCTCTGCTGCTCGCACCTAGAGCCCGGGAGAGGGGGGAGCGGAGCGTACGTGCTGC	352
QY	1983	CGCCGACCTGAGAGATGTGCTGTCTCCCTGATGTCCACAGTACTCATGTGCGTCACTG	2042
Db	353	CGCCGACCTGAGAGATGTGCTGTCTCCCTGATGTCCACAGTACTCATGTGCGTCACTG	412
QY	2043	CTGCCCCAGCTTCGTGCTCTCTGGGGAGAGGGCGCCCACTACGTGTGGCTGGGGGGTCTGG	2102
Db	413	CTGCCCCAGCTTCGTGCTCTCTGGGGAGAGGGCGCCCACTACGTGTGGCTGGGGGGTCTGG	472
QY	2103	CGGCTGCTCTCCGCTTCCTGCGCCCGCCCGCTTCTACCAAGAGGCTGAGAGCACTGGCTGTAC	2162
Db	473	CGGCTGCTCTCCGCTTCCTGCGCCCGCCCGCTTCTACCAAGAGGCTGAGAGCACTGGCTGTAC	532
QY	2163	TGCGCTTACACAGAGCATGTGTCTTCTTCTTTCGAGAAATTACACCGGGGCTTCAG	2216
Db	533	TGCGCTTACACAGAGCATGTGTCTTCTTCTTTCGAGAAATTACACCGGGGCTTCAG	586

RESULT 10	BE247552/c	340 bp	mRNA	linear	EST 03-OCT-2001
LOCUS	BE247552/c				
DEFINITION	BE247552				
	TCBAP2E4537				
	Patrol-HMSC				
	Project-TCBA				
	Homo sapiens				
	cdNA clone				
	TCBAP4537,				
	mRNA				
	sequence.				
ACCESSION	BE247552				
VERSION	BE247552.1				
KEYWORDS	GI:9099309				
SOURCE	EST.				
ORGANISM	Homo sapiens				
	Homo sapiens				
	(human)				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryotic: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo. 1 (Bases 1 to 340)	Wei, Y., Tsang, Y.T.M., Mei, G., Xu, J.M., Ali-Osman Jr., F.R., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.	Pediatric Leukemia cDNA Sequencing Project	Unpublished (2000)	Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@tcgcc.org
Citation: Carninci, P. and Hayashizaki, Y. High efficiency
full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)
Seq primer: M13 primer.

```

FEATURES
    source
        Location/Qualifiers
            1..340
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="TCBAP4537"
                /sex="male"
                /tissue_type="leukopheresis"
                /cell_type="pre-B cell"
                /dev_stage="pediatric 2 years"
                /lab_host="DH10B"
                /clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project=TCBA"
                /note="Vector: lambda PSB; Site 1: BamHI; Site 2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'-GGAGAGCTCGAGCGCGGAGGAGG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then digested. Second strand
was primed with a BamHI-dC primer
[5'-AGAGAGCTCGAGCTCGCGCGCGGCAATATATAT(C) 3']".
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda PSB vector. Library went through one round of

```


QY 1789 TCCGCGGAGCGGCGACGAGCCCGCACTCGCTTACCGCGGCGCGGCGGCGC 1848
DB 61 TCCGCGGAGCGGCGACGAGCCCGCACTCGCTTACCGCGGCGCGGCGGCGC 120
QY 1849 GCGCCATGCGGCTGGGCGGCGAGGCTGGGCGGCGGCGGCGGCGGCGGCGG 1908
DB 121 GCGCCATGCGGCTGGGCGGCGAGGCTGGGCGGCGGCGGCGGCGGCGGCGG 180
QY 1909 GCGCGGATGGCGGCTGGGCGGCGGCTGGGCGGCGGCGGCGGCGGCGG 1955
DB 181 GCGCGGATGGCGGCTGGGCGGCGGCTGGGCGGCGGCGGCGGCGGCGG 227

RESULT 13
BM789163 687 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST0068703 S11SNUI Homo sapiens cDNA clone S11SNUI-25-E07 5',
mRNA sequence.
ACCESSION BM789163
VERSION BM789163.1 GI:19137295
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 687)
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: Yongsung@mail.kribb.re.kr
Plate: 25 row: E column: 07
High quality sequence stop: 687.
Location/Qualifiers
1. 687
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S11SNUI-25-E07"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/cell_line="SNU-1"
/lab_host="Top10F"
/clone_lib="S11SNUI"
/note="Organ: Stomach, Vector: pME18-FU3, Site 1: XhoI;
Site 2: XhoI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deapped
with tabacco acid pyrophosphatase (TAP). The deapped
inact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18-FU3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match 9.6%; Score 223; DB 12; Length 687;
Best Local Similarity 100.0%; Pred. No. 1.2e-95;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1994 GAAGATGTGTGTCTCTGCTGCTGCTCAACAGCTACTTCATGCGCTACCTGTGCCAGCGT 2053
DB 1 GAAGATGTGTGTCTCTGCTGCTGCTGCTCAACAGCTACTTCATGCGCTACCTGTGCCAGCGT 60
QY 2054 CGTGTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2113
DB 61 CGTGTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 2114 GCGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2173
DB 121 GCGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 2174 GAGCATGTGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2216
DB 181 GAGCATGTGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 223

RESULT 14
BU931112 1216 bp mRNA linear EST 18-OCT-2002
LOCUS AGNCOURT 10475891 NIH MGC 109 Homo sapiens cDNA clone
DEFINITION IMAGE:6671220 5', mRNA sequence.
ACCESSION BU931112
VERSION BU931112.1 GI:24119931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1216)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Stransberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNU at:
http://image.llnl.gov
Plate: LNCM2949 row: m column: 12
High quality sequence stop: 269.
Location/Qualifiers
1. 1216
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6671220"
/tissue_type="teratocarcinoma, cell line"
/lab_host="VDH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/note="Organ: ovary; Vector: pOTB7, Site 1: EcoRI, Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGCGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

ORIGIN
Query Match 9.5%; Score 221; DB 13; Length 1216;
Best Local Similarity 100.0%; Pred. No. 1.1e-94;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1739 GTGAGCATCCCTGAGCATTCATCGGAGGAGGCGGCTTCCCTTGGCGCGCGG 1798
DB 34 GTGAGCATCCCTGAGCATTCATCGGAGGAGGCGGCTTCCCTTGGCGCGCGG 93
QY 1799 AGCGGCGAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1858

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 06:13:58 ; Search time 2582.32 Seconds
(without alignments)
12697.364 Million cell updates/sec

Title: US-09-901-484A-179_COPY_3899_4996

Perfect score: 1098

Sequence: 1 gtctctgtctccgcgattc.....agacttaagcttcgcat 1098

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2.1	209	13	BQ989814	OGF18L11
2	2.1	210	13	BQ847561	BQ847561 OGA3G09.Y
3	2.1	210	13	BQ857992	BQ857992 QGB9A22.Y
4	2.1	222	13	BQ992394	BQ992394 QGF25103.

5	2.1	223	13	BQ847396	BQ847396 OGA2A06.Y
6	2.1	259	13	BQ013709	BQ013709 QGJ5G07.Y
7	2.1	353	13	BQ981419	BQ981419 QGB13L15.
8	2.1	432	9	AV665533	AV665533 AV665533
9	2.1	551	13	BU011575	BU011575 QGJ16721.
10	2.1	607	9	AV665534	AV665534 AV665534
11	2.1	771	29	CNS02T0F	AL22508 Tetradon
12	2.1	921	29	CNS02T0F	AL21828 Tetradon
13	2.0	316	9	AV334716	AV334716 AV334716
14	2.0	367	28	A2726944	A2726944 RPT1-24-1
15	2.0	424	13	BQ597016	BQ597016 PERStoab2
16	2.0	466	9	AI431315	AI431315 ar55b09.x
17	2.0	498	28	BH742910	BH742910 g267b11.9
18	2.0	660	9	AI063012	AI063012 GH02422.5
19	2.0	663	28	BH450006	BH450006 BOMHJ6TF
20	2.0	668	29	CE279192	CE279192 ligR-g88-
21	2.0	682	14	CA993731	CA993731 PARI0A02
22	2.0	686	29	AG144022	AG144022 Pan treg1
23	2.0	726	28	B2508298	B2508298 BOMHJ6TF
24	2.0	786	28	B2070860	B2070860 IK179e10.
25	2.0	1757	11	AK013575	AK013575 Mus muscu
26	1.9	224	14	CP426180	CP426180 Iad12a11.
27	1.9	241	9	AV254417	AV254417 AV254417
28	1.9	261	9	AI032180	AI032180 or96a11.s
29	1.9	283	12	B1288613	B1288613 UI-R-DK0-
30	1.9	284	10	BB244365	BB244365 BB244365
31	1.9	292	9	AI435529	AI435529 th73e07.x
32	1.9	296	28	B2859441	B2859441 CH240_232
33	1.9	338	10	AW434852	AW434852 UI-R-B0P
34	1.9	368	13	BQ492779	BQ492779 EST01945
35	1.9	430	12	BM106873	BM106873 510932 MA
36	1.9	436	29	CE610771	CE610771 ligR-g88-
37	1.9	475	21	CC557326	CC557326 CH240_465
38	1.9	481	13	BX721433	BX721433 BX721433
39	1.9	492	28	B2845349	B2845349 CH240_250
40	1.9	516	12	B1815360	B1815360 PIRStoaa1
41	1.9	536	13	BQ521596	BQ521596 NISC n111
42	1.9	544	14	W52378	W52378 zc48c05.g1
43	1.9	545	28	BH930122	BH930122 c0198a04.
44	1.9	579	12	BU542168	BU542168 BU542168
45	1.9	581	12	BI586329	BI586329 RH26255.5

ALIGNMENTS

RESULT 1	BQ989814	209 bp	mrna	linear	EST 21-AUG-2002
LOCUS	OGF18L11	yg abl OG_EFGHJ	lettuce	seiriola	Lactuca sativa cDNA clone
DEFINITION	OGF18L11	yg abl OG_EFGHJ	lettuce	seiriola	Lactuca sativa cDNA clone
ACCESSION	BQ989814				
VERSION	BQ989814.1	GI:22409349			
KEYWORDS	EST				
SOURCE	Lactuca sativa				
ORGANISM	Lactuca sativa				
REFERENCE					
AUTHORS	Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Lai, Z., Church, S., Jackson, L. and Bradford, K.				
TITLE	Lettuce and Sunflower ESTs from the Compositae Genome Project				
COMMENT	http://compgenome.ucdavis.edu/unpublished (2002)				
JOURNAL	Contact: Alexander Kozik [R.W.Michelmore]				
COMMENT	Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Aasmundson Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742				

Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QG_CA_Contig7771, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: QGF18 row: L column: 11.
 Location/Qualifiers
 1..209

/organism="Lactuca sativa"
 /mol_type="mRNA"
 /cultivar="L.seriola"
 /db_xref="taxon:4236"
 /clone="QG18L11"
 /lab_host="E.coli"
 /clone_lib="QG_EFGHJ lettuce seriola"
 /note="Vector: pBRCDNASfiAB. The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>
 TAG_TISSUE=roots
 TAG_LIB=QG_EFGHJ lettuce seriola
 TAG_SEQ=GTTCACGGG"

ORIGIN

Query Match 2.1%; Score 23; DB 13; Length 209;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 480 TTTATTAATTATATATATTTAA 502
 ||||||||||||||||||||
 Db 174 TTTATTAATTATATATATTTAA 196

RESULT 2
 BQ847561 210 bp mRNA linear EST 14-ANG-2002
 LOCUS QGA3G09.yg.ab1 QG ABCDI lettuce salinas Lactuca sativa cDNA clone
 DEFINITION BQ847561
 ACCESSION BQ847561
 VERSION BQ847561.1 GI:22231980
 KEYWORDS EST.
 SOURCE Lactuca sativa
 ORGANISM Lactuca sativa
 Lactuca sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asterales; Cichorioideae; Cichorioideae; Lactuca.
 1 (bases 1 to 210)
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Composite Genome Project
<http://compgenomics.ucdavis.edu/>
 Unpublished (2002)
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QG_CA_Contig7771, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: QGA3 row: 9 column: 09.
 Location/Qualifiers
 1..210
 /organism="Lactuca sativa"
 /mol_type="mRNA"

FEATURES

source
 1..210
 /organism="Lactuca sativa"
 /mol_type="mRNA"

ORIGIN

Query Match 2.1%; Score 23; DB 13; Length 210;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 480 TTTATTAATTATATATTTAA 502
 ||||||||||||||||||||
 Db 174 TTTATTAATTATATATTTAA 196

RESULT 3
 BQ857992 210 bp mRNA linear EST 14-AUG-2002
 LOCUS QGB9A22.yg.ab1 QG ABCDI lettuce salinas Lactuca sativa cDNA clone
 DEFINITION QGB9A22, mRNA sequence.
 ACCESSION BQ857992
 VERSION BQ857992.1 GI:22243457
 KEYWORDS EST.
 SOURCE Lactuca sativa
 ORGANISM Lactuca sativa
 Lactuca sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asterales; Cichorioideae; Cichorioideae; Lactuca.
 1 (bases 1 to 210)
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Composite Genome Project
<http://compgenomics.ucdavis.edu/>
 Unpublished (2002)
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QG_CA_Contig7771, see <http://cgpdb.ucdavis.edu/>
 for details.
 Plate: QGB9 row: A column: 22.
 Location/Qualifiers
 1..210
 /organism="Lactuca sativa"
 /mol_type="mRNA"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="QGB9A22"
 /lab_host="E.coli"
 /clone_lib="QG ABCDI lettuce salinas"
 /note="Vector: pBRCDNASfiAB. The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each

FEATURES

source

source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpb.ucdavis.edu/>
TAG LIB=OG ABCDI lettuce salinas
TAG_SEQ=GTGCACGGG"

Query Match 2.1%; Score 23; DB 13; Length 210;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 480 TTTATTAATTTATATATTTAA 502
|||||
174 TTTATTAATTTATATATTTAA 196

RESULT 4
LOCUS B0992394 222 bp mRNA linear EST 21-AUG-2002
DEFINITION OG25103.yg.ab1 OG_EFGHU lettuce serriola lactuca sativa cDNA clone
OG25103, mRNA sequence.

ACCESSION B0992394
VERSION B0992394.1 GI:22411929
KEYWORDS EST.

SOURCE Lactuca sativa

ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.

REFERENCE 1 (bases 1 to 222)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://compenomics.ucdavis.edu/>
Unpublished (2002)

JOURNAL Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659

COMMENT Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OG_Ca_Contig7771, see <http://cgpb.ucdavis.edu/>
for details.
Plate: OG25 row: 1 column: 03.

FEATURES
source Location/Qualifiers

1..222
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="OG25103"
/lab_host="E.coli"
/clone_lib="OG ABCDI lettuce serriola"
/note="Vector: pBRCDNASFIAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpb.ucdavis.edu/>
TAG LIB=OG EFGHU lettuce serriola
TAG_SEQ=GTGCACGGG"

ORIGIN

Query Match 2.1%; Score 23; DB 13; Length 222;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 480 TTTATTAATTTATATATTTAA 502
|||||
174 TTTATTAATTTATATATTTAA 196

RESULT 5
LOCUS B0847396

DEFINITION OG2A06.yg.ab1 OG ABCDI lettuce salinas Lactuca sativa cDNA clone
OG2A06, mRNA sequence.

ACCESSION B0847396
VERSION B0847396.1 GI:22231616
KEYWORDS EST.

SOURCE Lactuca sativa

ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.

REFERENCE 1 (bases 1 to 223)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://compenomics.ucdavis.edu/>
Unpublished (2002)

JOURNAL Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659

COMMENT Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OG_Ca_Contig7771, see <http://cgpb.ucdavis.edu/>
for details.
Plate: OG2A row: a column: 06.

FEATURES
source Location/Qualifiers

1..223
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="OG2A06"
/lab_host="E.coli"
/clone_lib="OG ABCDI lettuce salinas"
/note="Vector: pBRCDNASFIAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpb.ucdavis.edu/>
TAG LIB=OG ABCDI lettuce salinas
TAG_SEQ=GTGCACGGG"

ORIGIN

Query Match 2.1%; Score 23; DB 13; Length 223;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 480 TTTATTAATTTATATATTTAA 502
|||||
174 TTTATTAATTTATATATTTAA 196

ORIGIN

RESULT 6
BU013709
LOCUS
DEFINITION BU013709 259 bp mRNA linear EST 22-AUG-2002
Q6J5G07.yg.ab1 OG_EFGHU lettuce serriola Lactuca sativa cDNA clone
Q6J5G07, mRNA sequence.
ACCESSION
VERSION BU013709
KEYWORDS
SOURCE Lactuca sativa
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 259)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compenomics.ucdavis.edu/
Unpublished (2002)
JOURNAL
COMMENT
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmudson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OG_CA_Config7771, see http://cgpdb.ucdavis.edu/
for details.
Plate: Q6J5 row: G column: 07.

FEATURES
source
location/Qualifiers
1..259
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="Q6J5G07"
/lab_host="E.coli"
/clone_lib="OG_EFGHU lettuce serriola"
/note="Vector: pBRCDNA5flab; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG TISSUE=roots
TAG LIB=OG_EFGHU lettuce serriola
TAG_SEQ=GTTCACGCGG"

ORIGIN
Query Match 2.1%; Score 23; DB 13; Length 259;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 480 TTTATTATTATATATATATAA 502
|||||
Db 175 TTTATTATTATATATATATAA 197
|||||

RESULT 7
BQ981419
LOCUS
DEFINITION BQ981419 353 bp mRNA linear EST 21-AUG-2002
QGE13115.yg.ab1 OG_EFGHU lettuce serriola Lactuca sativa cDNA clone
QGE13115, mRNA sequence.
ACCESSION
VERSION BQ981419
KEYWORDS
SOURCE Lactuca sativa
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 432)
Takasuga, A., Hirotsune, S., Itoh, R., Jitchazon, A., Suzuki, H., Aso, H.,
and Sugimoto, Y.
Establishment of a high throughput EST sequencing system using

SOURCE
ORGANISM
Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichoriaceae; Lactuca.
1 (bases 1 to 353)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compenomics.ucdavis.edu/
Unpublished (2002)
JOURNAL
COMMENT
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmudson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OG_CA_Config7771, see http://cgpdb.ucdavis.edu/
for details.
Plate: QGE13 row: L column: 15.

FEATURES
source
location/Qualifiers
1..353
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGE13115"
/lab_host="E.coli"
/clone_lib="OG_EFGHU lettuce serriola"
/note="Vector: pBRCDNA5flab; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG TISSUE=flowers environmental stress
TAG LIB=OG_EFGHU lettuce serriola
TAG_SEQ=CGAATCGCGG"

ORIGIN
Query Match 2.1%; Score 23; DB 13; Length 353;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 480 TTTATTATTATATATATATAA 502
|||||
Db 104 TTTATTATTATATATATATAA 126
|||||

RESULT 8
AV665533
LOCUS
DEFINITION AV665533 432 bp mRNA linear EST 28-NOV-2001
AV665533 Bos taurus brain fetus Bos taurus cDNA clone E1BR049B04
3', mRNA sequence.
ACCESSION
VERSION AV665533
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 432)
Takasuga, A., Hirotsune, S., Itoh, R., Jitchazon, A., Suzuki, H., Aso, H.,
and Sugimoto, Y.
Establishment of a high throughput EST sequencing system using

poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
JOURNAL
MEDLINE
PUBMED
21570554
11713328
COMMENT
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@ccoca.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers
1..432

FEATURES
source
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR049B04"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus brain fetus"
/note="Vector: pZLI; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"

ORIGIN

Query Match 2.1%; Score 23; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 480 TTTATTATTATATATATTTAA 502
|||||
Db 350 TTTATTATTATATATTTAA 372

RESULT 9
BU011575 551 bp mRNA linear EST 22-AUG-2002
LOCUS
OCG16J21.YG.ab1 OG_EFGH lettuce serritola lactuca sativa cDNA clone
DEFINITION
OCG16J21, mRNA sequence.
ACCESSION
BU011575
VERSION
BU011575.1 GI:22445970
KEYWORDS
EST.
SOURCE
Lactuca sativa
ORGANISM
Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.
1 (bases 1 to 551)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L.,
Lin,H., Van Damme,M., Lavelle,D., Chevaller,P., Ziegler,J.,
Ellison,P., Koltman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compgenomics.ucdavis.edu/
Unpublished (2002)
CONTACT: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Assumedeon Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OG_CA_Contig777, see http://cgpdb.ucdavis.edu/
for details.
Plate: OCG16 row: J column: 21.

FEATURES
source
Location/Qualifiers
1..551
/organism="Lactuca sativa"
/mol_type="mRNA"

/cultivar="L. serritola"
/db_xref="taxon:4236"
/clone="OCG16J21"
/lab_host="HE.coli"
/clone_lib="OG_EFGH lettuce serritola"
/note="Vector: pARCDNA51AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=flowers pre-fertilized
TAG_LIB=OG_EFGH lettuce serritola
TAG_SEQ=GCTTACGCG3"

ORIGIN

Query Match 2.1%; Score 23; DB 13; Length 551;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 480 TTTATTATTATATATTTAA 502
|||||
Db 372 TTTATTATTATATATTTAA 394

RESULT 10
AV665534 607 bp mRNA linear EST 28-NOV-2001
LOCUS
AV665534 Bos taurus brain fetus Bos taurus cDNA clone E1BR049B04
DEFINITION
AV665534 5', mRNA sequence.
ACCESSION
AV665534
VERSION
AV665534.1 GI:3924564
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 607)
Takasuga,A., Hitotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aeo,H.
and Sugimoto,Y.
Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
JOURNAL
MEDLINE
PUBMED
21570554
11713328
COMMENT
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@ccoca.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers
1..607

FEATURES
source
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR049B04"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus brain fetus"
/note="Vector: pZLI; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"

ORIGIN

Query Match 2.1%; Score 23; DB 9; Length 607;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 480 TTTATTAAATTTATATATATTA 502
|||||
451 TTTATTAAATTTATATATATTA 429

Db

RESULT 11
CNS0300J/c 771 bp DNA linear GSS 01-SEP-2000
LOCUS Tetradon nigroviridis genome survey sequence PUC-Orl end of clone
DEFINITION 203123 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL222508.1 GI:7881327
VERSION GSS; genome survey sequence.
KEYWORDS Tetradon nigroviridis
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetradon.

REFERENCE
AUTHORS Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645

REFERENCE
AUTHORS 2
Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143

REFERENCE
AUTHORS 3 (bases 1 to 771)
Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.
location/Qualifiers

FEATURES
source 1..771
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="203123"
/clone_lib="G"
/note="Genoscope sequence ID : COAG203AE12SP1-end :
PUC-Orl"

ORIGIN
Query Match 2.1%; Score 23; DB 29; Length 771;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 TCTTTTCTTGAATGTTTT 363
|||||
545 TCTTTTCTTGAATGTTTT 523

Db

RESULT 12

CNS02TZP 921 bp DNA linear GSS 01-SEP-2000
LOCUS Tetradon nigroviridis genome survey sequence T7 end of clone
DEFINITION 166C04 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL213828.1 GI:7872647
VERSION GSS; genome survey sequence.
KEYWORDS Tetradon nigroviridis
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetradon.

REFERENCE
AUTHORS 1
Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645

REFERENCE
AUTHORS 2
Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143

REFERENCE
AUTHORS 3 (bases 1 to 921)
Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.
location/Qualifiers

FEATURES
source 1..921
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="166C04"
/clone_lib="G"
/note="Genoscope sequence ID : COAG166B02LP1-end : T7"

ORIGIN
Query Match 2.1%; Score 23; DB 29; Length 921;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 TCTTTTCTTGAATGTTTT 363
|||||
697 TCTTTTCTTGAATGTTTT 719

Db

RESULT 13
AV334716 236 bp mRNA linear EST 11-NOV-1999
LOCUS AV334716 RIKEN full-length enriched, adult male medulla oblongata
DEFINITION Mus musculus cDNA clone 6330563N19 3', mRNA sequence.
ACCESSION AV334716.1 GI:6374768
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

SOURCE Plasmodium falciparum 3D7
 ORGANISM Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 (bases 1 to 424)
 AUTHORS Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Page, D.,
 Marra, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I.,
 Tsagaris, H., R., Belaygorod, L., Franklin, C., Carr, L., Grow, A.,
 Maguire, L., Richey, J., Watkins, J., Kennedy, S., Levine, D.,
 Waterston, R., Wilson, R. and Sibley, D.
 TITLE Washu Plasmodium EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: L. David Sibley
 Washu Plasmodium EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watscn.wustl.edu
 Library was constructed by Debopam Chakrabarti DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: L. David Sibley
 (sibley@borcim.wustl.edu), Washington University
 Seq primer: -40UP from Gibco
 High quality sequence stop: 423.
 Location/Qualifiers

FEATURES

source
 1..424
 /organism="Plasmodium falciparum 3D7"
 /mol_type="mRNA"
 /db_xref="taxon:36329"
 /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
 /clone_lib="Plasmodium falciparum 3D7 asexual cDNA"
 /note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:
 XhoI; Library was constructed by Debopam Chakrabarti.
 Total RNA samples were isolated from mixed stage
 saponin(0.1%) lysed P. falciparum 3D7 infected
 erythrocytes by the acidic guanidinium-phenol chloroform
 method. The poly A+ RNA was isolated by the polyAT-Tract
 mRNA isolation system (Promega, WI) using streptavidin
 Magneticbead particles. Directional cDNA libraries were
 constructed by oligo d(T) priming of poly(A)+ RNA (5mg)
 into EcoRI and XhoI sites of 1 ZapII vector using the Zap
 cDNA synthesis kit (Stratagene, CA). The average size of the
 cDNA inserts in the library was between 1.0 and 1.5kb.
 Clones were mass excised using the ExAssist helper phage
 (Stratagene), the phagemids were precipitated with PEG
 8000 and extracted with phenol/chloroform. Phagemid DNA
 was electroporated into DH10B cells."

ORIGIN

Query Match 2.0%; Score 22; DB 13; Length 424;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 TTTAAATATAAATTTTAA 531
 |||||
 Db 127 TTTAAATATAAATTTTAA 148
 |||||

Search completed: May 26, 2004, 12:51:08
 Job time : 2588.32 secs